

AROUND NUMBERS: COLLOCATIONAL PATTERNS INVOLVING NUMBERS IN BUSINESS ENGLISH

Combinations of numbers and approximators

Joint 10th ABC Europe
convention + 2nd GABC
conference
27-29 May 2010
Antwerp, Belgium

Diane Goossens*
Sylvie De Cock*
Philippe Hilgsmann

Université catholique de Louvain
*Centre for English Corpus Linguistics

AROUND NUMBERS

- ▶ Research project: ‘quantification linguistic toolbox’ in business language
 - = linguistic devices used to express quantity and approximate quantity
- ▶ Exploratory study
 - Focus on numbers
 - Collocational patterns involving approximating devices and numbers
 - Computerized corpora of Business English:
 - Academic Business English
 - Business News Reporting

AROUND NUMBERS

- ▶ Link between quantification, approximation and vague language
- ▶ Channell 1994 ‘Vague Language’
- ▶ Corpus-based studies
 - Jucker et al. (2003); Ruzaitè (2007); Koester (2007)
- ▶ Make use of pre-established lists of approximators
 - Limited:
 - *about, around*
 - *approximately, roughly, round, like and or so*



METHOD

4

METHOD

- Corpus-driven method (Tognini-Bonelli 2001): using corpus linguistics as a discovery procedure
- Starting point: numbers (e.g. *15, hundred, three quarters*)
- Identify any device used around numbers to approximate a quantity
- POS-tagged corpora (CLAWS 7)
 - 10 tags for numbers (e.g. MC or MD) automatically retrieved using Concord in WordSmith Tools (Scott 2004)

METHOD

File Edit View Compute Settings Windows Help

N	Concordance	et	ag	ord #	#	s	#	s	#	s	#	s	File	%
1	price_NN1 rises_VVZ a_AT1 key_JJ priority_NN1 for_IF 2007_MC _.	0.647	91	%	0	%	0	%	0	%	0	%	os tagged.bt	28%
2	over_RP as_CSA Microsoft_NP1 CEO_NN1 in_II 2000_MC _.	6.005	86	%	0	%	0	%	0	%	0	%	os tagged.bt	45%
3	capable_JJ handsets_NN2 reaching_VVG 50m_NNU in_II 2009_MC _.	3.817	27	%	0	%	0	%	0	%	0	%	os tagged.bt	30%
4	fell_VVD 2.8_MC per_NNU21 cent_NNU22 to_II 141_MC _.	1.602	26	%	0	%	0	%	0	%	0	%	os tagged.bt	28%
5	down_RP 4.9_MC per_NNU21 cent_NNU22 at_II 68.45_MC _.	7.393	92	%	0	%	0	%	0	%	0	%	os tagged.bt	42%
6	2_MC per_NNU21 cent_NNU22 higher_JJR at_II 22.29_MC _.	9.381	37	%	0	%	0	%	0	%	0	%	os tagged.bt	44%
7	outweighs_VVZ the_AT [_(drawbacks_NN2])_MC	0.973	53	%	0	%	0	%	0	%	0	%	os tagged.bt	41%
8	The_AT Financial_JJ Times_NNT2 Limited_JJ 2006_MC	7.456	82	%	0	%	0	%	0	%	0	%	os tagged.bt	26%
9	placed_VVN a_AT1 50p_NNU accumulator_NN1 on_II eight_MC horses_NN2 and_CC won_VVN 1million_FO after_CS	6.190	21	%	0	%	0	%	0	%	0	%	os tagged.bt	93%
10	_, especially_RR in_II Delhi_NP1 _, where_RRQ only_RR two_MC slots_NN2 are_VBR on_II offer_NN1 _ In_II some_DD	3.916	25	%	0	%	0	%	0	%	0	%	os tagged.bt	9%
11	kisses_NN2 Heard_VVD this_DD1 one_PN1 ?_? Two_MC women_NN2 are_VBR walking_VVG through_II the_AT	4.893	39	%	0	%	0	%	0	%	0	%	os tagged.bt	63%
12	_. The_AT figures_NN2 are_VBR based_VVN on_II 10_MC analyst_NN1 forecasts_NN2 collated_VVD by_II Standard_JJ	0.131	58	%	0	%	0	%	0	%	0	%	os tagged.bt	97%
13	currently_RR operates_VVZ 33_MC aircraft_NN on_II 43_MC short-haul_JJ routes_NN2 with_IVW fewer_DAR daily_JJ	5.737	10	%	0	%	0	%	0	%	0	%	os tagged.bt	44%
14	it_PPH1 would_VM deliver_VVI in_II total_NN1 only_RR 9_MC A380s_FO rather_II21 than_II22 the_AT 25_MC	6.441	57	%	0	%	0	%	0	%	0	%	os tagged.bt	42%
15	and_CC stamina_NN1 from_II which_DDQ only_RR two_MC outcomes_NN2 are_VBR possible_JJ _ Either_RR Mr_NNB	4.601	46	%	0	%	0	%	0	%	0	%	os tagged.bt	20%
16	"_ For_IF a_AT1 small_JJ employer_NN1 with_IVW only_RR five_MC to_II ten_MC staff_NN _, it_PPH1 is_VBZ very_RG	4.983	46	%	0	%	0	%	0	%	0	%	os tagged.bt	76%
17	mergers_NN2 and_CC acquisitions_NN2 _, with_IVW only_RR 7_MC per_NNU21 cent_NNU22 saying_VVG that_CST they_PPHS2	9.410	02	%	0	%	0	%	0	%	0	%	os tagged.bt	99%
18	months_NNT2 of_IO last_MD year_NNT1 _, with_IVW only_RR 9_MC per_NNU21 cent_NNU22 reporting_VVG a_AT1 rise_NN1 _.	1.385	54	%	0	%	0	%	0	%	0	%	os tagged.bt	90%
19	free_VVI it_PPH1 from_II their_APPGE onerous_JJ 12_MC per_NNU21 cent_NNU22 interest_NN1 charge_NN1 _.	5.496	70	%	0	%	0	%	0	%	0	%	os tagged.bt	68%
20	"_ For_IF a_AT1 small_JJ employer_NN1 with_IVW only_RR five_MC to_II ten_MC staff_NN _, it_PPH1 is_VBZ very_RG	9.947	56	%	0	%	0	%	0	%	0	%	os tagged.bt	76%
21	and_CC has_VHZ been_VBN working_VVG only_RR two_MC days_NNT2 a_AT1 week_NNT1 since_II September_NPM1	8.784	17	%	0	%	0	%	0	%	0	%	os tagged.bt	93%
22	separate_JJ options_NN2 grants_NN2 made_VVN on_II 42_MC dates_NN2 _ Backdating_NP1 involves_VVZ setting_VVG	7.157	69	%	0	%	0	%	0	%	0	%	os tagged.bt	26%
23	Tories_NN2 a_AT1 lead_NN1 of_IO 13_MC points_NN2 on_II 45_MC per_NNU21 cent_NNU22 _, with_IVW Labour_JJ on_II 32_MC	3.613	92	%	0	%	0	%	0	%	0	%	os tagged.bt	68%
24	instead_RR _ Mr_NNB Ellison_NP1 is_VBZ keen_JJ on_II two_MC bits_NN2 of_IO Suns_NN2 software_NN1 portfolio_NN1	1.369	03	%	0	%	0	%	0	%	0	%	os tagged.bt	19%
25	per_NNU21 cent_NNU22 _, with_IVW Labour_JJ on_II 32_MC per_NNU21 cent_NNU22 _ Support_VVO for_IF the_AT	3.620	92	%	0	%	0	%	0	%	0	%	os tagged.bt	68%
26	_ Mr_NNB Albanese_NP1 reckons_VVZ on_II two_MC difficult_JJ years_NNT2 before_II commodity_NN1	9.117	19	%	0	%	0	%	0	%	0	%	os tagged.bt	9%
27	the_AT banks_NN2 resources_NN2 on_II two_MC of_IO the_AT riskiest_JJT areas_NN2 of_IO the_AT	7.914	93	%	0	%	0	%	0	%	0	%	os tagged.bt	97%
28	it_PPH1 would_VM announce_VVI results_NN2 on_II 2_MC March_NPM1 _ HSBC_NP1 has_VHZ not_XX ruled_VVN	4.691	99	%	0	%	0	%	0	%	0	%	os tagged.bt	69%
29	eight_MC in_II the_AT series_NN _, released_VVN on_II 18_MC November_NPM1 _, but_CCB the_AT game_NN1 met_VVD	9.027	12	%	0	%	0	%	0	%	0	%	os tagged.bt	58%
30	which_DDQ have_VH0 left_VVN it_PPH1 reliant_JJ on_II two_MC cash_NN1 infusions_NN2 from_II the_AT US_NP1	9.182	16	%	0	%	0	%	0	%	0	%	os tagged.bt	58%
31	sales_NN fell_VVD 23_MC per_NNU21 cent_NNU22 on_II 2007_MC _ Ford_NP1 said_VVD yesterday_RT that_CST sales_NN	1.788	39	%	0	%	0	%	0	%	0	%	os tagged.bt	78%
32	up_RP about_RG 24_MC per_NNU21 cent_NNU22 on_II 2005_MC despite_II Vodafones_NP2 14.9bn_FO loss_NN1 for_IF	1.823	35	%	0	%	0	%	0	%	0	%	os tagged.bt	45%
33	by_II some_DD 75_MC per_NNU21 cent_NNU22 on_II 2008_MC _ which_DDQ itself_PPX1 saw_VVD a_AT1 reduction_NN1	2.171	39	%	0	%	0	%	0	%	0	%	os tagged.bt	53%
34	down_RP 1.6_MC per_NNU21 cent_NNU22 on_II 2007_MC http_NNU _ : //www.timesonline.co.uk/	0.870	90	%	0	%	0	%	0	%	0	%	os tagged.bt	76%
35	_, up_II 14_MC per_NNU21 cent_NNU22 on_II 2007_MC _ according_II21 to_II22 the_AT latest_JJT figures_NN2 _.	9.419	00	%	0	%	0	%	0	%	0	%	os tagged.bt	63%
36	two-year_JJ one_PN1 instead_RR _ focusing_VVG on_II two_MC things_NN2 _ internal_JJ reorganisation_NN1 and_CC	1.513	96	%	0	%	0	%	0	%	0	%	os tagged.bt	14%
37	operating_NN1 system_NN1 -_ installed_VVN on_II nine_MC out_II21 of_II22 ten_MC personal_JJ computers_NN2 -_	1.451	04	%	0	%	0	%	0	%	0	%	os tagged.bt	50%
38	_, such_II21 as_II22 service_NN1 charges_NN2 _ on_II 70_MC empty_JJ stores_NN2 _ Market_VVO sources_NN2	9.189	25	%	0	%	0	%	0	%	0	%	os tagged.bt	53%

concordance collocates plot patterns clusters filenames source text notes

9.023 | Set | tion_NN1 rests_VVZ largely_RR on_II the_AT efforts_NN2 of_IO one_MC1 or_CC two_MC companies_NN2 _ such_II21 as_II22 Infosys_NP1 _ which_DDQ are_VBR impecc

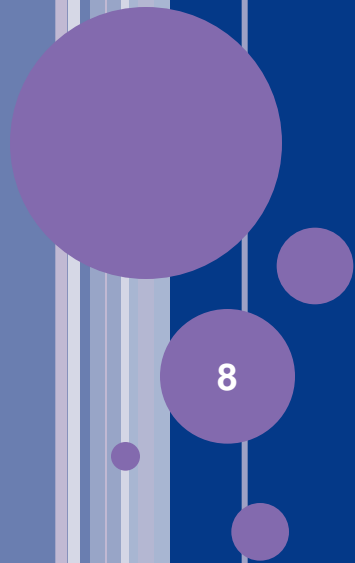
METHOD

File Edit View Compute Settings Windows Help

N	Concordance	et	ag	ord #	#	s	#	s	#	s	#	s	File	%
297	at_II like-for-like_JJ growth_NN1 of_IO about_RG 7_MC per_NNU21 cent_NNU22 last_MD time_NNT1 and_CC	6.700	68	%	0	%	0	%	0	%	0	%	os tagged.txt	42%
298	of_IO grey_JJ at_II a_AT1 resolution_NN1 of_IO about_RG 170_MC pixels_NN2 per_II inch_NNU1 ., which_DDQ is_VBZ	8.040	01	%	0	%	0	%	0	%	0	%	os tagged.txt	26%
299	operated_VVN only_RR about_RG 17_MC of_IO about_RG 500_MC routes_NN2 ., While_CS there_EX are_VBR few_DA2	8.808	12	%	0	%	0	%	0	%	0	%	os tagged.txt	31%
300	exit_NN1 price/earnings_FU ratio_NN1 of_IO about_RG 20_MC times_NNT2 next_MD years_NNT2 estimated_VVD	7.575	11	%	0	%	0	%	0	%	0	%	os tagged.txt	38%
301	tracker_NN1 loan_NN1 is_VBZ now_RT about_RG 380_MC lower_JJR than_CSN in_II January_NPM1 2008_MC .,	4.141	12	%	0	%	0	%	0	%	0	%	os tagged.txt	76%
302	., " " The_AT pound_NN1 is_VBZ now_RT about_RG 30_MC per_NNU21 cent_NNU22 below_II its_APPGE peak_NN1 .,	7.654	79	%	0	%	0	%	0	%	0	%	os tagged.txt	61%
303	this_DD1 ., " " The_AT ABI_NP1 monitors_VVZ about_RG 700_MC companies_NN2 and_CC only_JJ issues_NN2 red-top_JJ	1.323	29	%	0	%	0	%	0	%	0	%	os tagged.txt	65%
304	company_NN1 ., which_DDQ makes_VVZ about_RG 70_MC per_NNU21 cent_NNU22 of_IO its_APPGE profit_NN1 from_II	7.639	14	%	0	%	0	%	0	%	0	%	os tagged.txt	38%
305	150,000_MC tracker_NN1 loan_NN1 is_VBZ about_RG 375_MC lower_JJR than_CSN in_II October_NPM1 last_MD	2.549	98	%	0	%	0	%	0	%	0	%	os tagged.txt	80%
306	number_NN1 of_IO unemployed_JJ is_VBZ about_RG 50_MC per_NNU21 cent_NNU22 higher_JJR than_CSN a_AT1	2.068	54	%	0	%	0	%	0	%	0	%	os tagged.txt	84%
307	past_JJ nine_MC months_NNT2 involving_VVG about_RG 200_MC of_IO its_APPGE staff_NN from_II different_JJ areas_NN2	4.301	40	%	0	%	0	%	0	%	0	%	os tagged.txt	9%
308	council_NN1 is_VBZ not_XX immune_JJ ., About_RG 300_MC council_NN1 jobs_NN2 are_VBR expected_VVN to_TO	3.968	06	%	0	%	0	%	0	%	0	%	os tagged.txt	93%
309	1,427_MC employees_NN2 ., However_RR ., about_RG 300_MC staff_NN - - of_IO whom_PNQO the_AT majority_NN1	4.840	07	%	0	%	0	%	0	%	0	%	os tagged.txt	47%
310	1,427_MC employees_NN2 ., However_RR ., about_RG 300_MC staff_NN - - of_IO whom_PNQO the_AT majority_NN1	9.250	64	%	0	%	0	%	0	%	0	%	os tagged.txt	48%
311	1,427_MC employees_NN2 ., However_RR ., about_RG 300_MC staff_NN - - of_IO whom_PNQO the_AT majority_NN1	5.360	30	%	0	%	0	%	0	%	0	%	os tagged.txt	47%
312	ownership_NN1 plan_NN1 holding_VVG about_RG 10_MC per_NNU21 cent_NNU22 ., Ryanair_NP1 said_VVD the_AT	8.760	11	%	0	%	0	%	0	%	0	%	os tagged.txt	31%
313	., Savers_NN2 currently_RR have_VH0 about_RG 944_MC billion_NNO held_VVN in_II banks_NN2 and_CC	7.626	81	%	0	%	0	%	0	%	0	%	os tagged.txt	81%
314	the_AT UK_NP1 ., where_CS it_PPH1 has_VHZ about_RG 240_MC managed_VVD and_CC franchised_VVD properties_NN2 .,	8.493	63	%	0	%	0	%	0	%	0	%	os tagged.txt	79%
315	to_TO fall_VVI ., It_PPH1 has_VHZ about_RG 8_MC per_NNU21 cent_NNU22 of_IO US_NP1 searches_NN2	5.344	84	%	0	%	0	%	0	%	0	%	os tagged.txt	22%
316	., In_II total_NN1 ., the_AT site_NN1 has_VHZ about_RG 700_MC pieces_NN2 of_IO content_NN1 ranging_VVG from_II	1.405	29	%	0	%	0	%	0	%	0	%	os tagged.txt	27%
317	previous_JJ years_NNT2 weve_NN1 had_VHD about_RG 180_MC to_II 200_MC people_NN ., We_PPIS2 had_VHD 280_MC	0.510	03	%	0	%	0	%	0	%	0	%	os tagged.txt	94%
318	the_AT sector_NN1 and_CC guaranteed_VVN about_RG 100_MC billion_NNO out_II21 of_II22 a_AT1 proposed_JJ 250_MC	8.623	22	%	0	%	0	%	0	%	0	%	os tagged.txt	83%
319	about_RG 5bn_NNU ., having_VHG grown_VVN about_RG 10_MC per_NNU21 cent_NNU22 a_AT1 year_NNT1 over_II the_AT	9.076	27	%	0	%	0	%	0	%	0	%	os tagged.txt	44%
320	is_VBZ forecast_VVN to_TO grow_VVI about_RG 50_MC per_NNU21 cent_NNU22 to_II more_DAR than_CSN	9.444	51	%	0	%	0	%	0	%	0	%	os tagged.txt	28%
321	in_II the_AT US_NP1 ., It_PPH1 generates_VVZ about_RG 15_MC per_NNU21 cent_NNU22 of_IO its_APPGE earnings_NN2	2.949	76	%	0	%	0	%	0	%	0	%	os tagged.txt	34%
322	said_VVD Goldman_NP1 generated_VVD about_RG 60_MC per_NNU21 cent_NNU22 of_IO its_APPGE revenue_NN1	3.032	26	%	0	%	0	%	0	%	0	%	os tagged.txt	41%
323	FTSE_NN1 Eurofirst_NN1 300_MC gained_VVD about_RG 30_MC per_NNU21 cent_NNU22 in_II dollar_NNU1 terms_NN2 in_II	8.934	33	%	0	%	0	%	0	%	0	%	os tagged.txt	28%
324	within_II its_APPGE existing_JJ funds_NN2 ., About_RG 42_MC per_NNU21 cent_NNU22 of_IO Polar_JJ 's_GE funds_NN2	7.779	76	%	0	%	0	%	0	%	0	%	os tagged.txt	35%
325	per_NNU21 cent_NNU22 stake_NN1 ., From_II about_RG 4_MC three_MC years_NNT2 ago_RA ., the_AT shares_NN2	5.394	73	%	0	%	0	%	0	%	0	%	os tagged.txt	34%
326	increase_VVI their_APPGE leverage_NN1 from_II about_RG 15_MC to_II one_MC1 to_II as_RG much_DA1 as_CSA 40_MC to_II	7.002	58	%	0	%	0	%	0	%	0	%	os tagged.txt	81%
327	in_II the_AT City_NN1 - - rose_VVD from_II about_RG 300_MC to_II 900_MC ., " " Only_RR the_AT management_NN1	1.350	06	%	0	%	0	%	0	%	0	%	os tagged.txt	76%
328	index_NN1 ., which_DDQ accounts_VVZ for_IF about_RG 12_MC per_NNU21 cent_NNU22 of_IO the_AT US_NP1	8.598	05	%	0	%	0	%	0	%	0	%	os tagged.txt	75%
329	manufacturing_NN1 accounts_VVZ for_IF about_RG 20_MC per_NNU21 cent_NNU22 of_IO eurozone_NN1 GDP_NN1 .,	8.794	35	%	0	%	0	%	0	%	0	%	os tagged.txt	65%
330	eMusics_NN1 service_NN1 account_NN1 for_IF about_RG 30_MC per_NNU21 cent_NNU22 of_IO the_AT market_NN1 .,	2.029	87	%	0	%	0	%	0	%	0	%	os tagged.txt	41%
331	., which_DDQ accounts_VVZ for_IF about_RG 40_MC per_NNU21 cent_NNU22 of_IO economic_JJ growth_NN1	8.274	03	%	0	%	0	%	0	%	0	%	os tagged.txt	44%
332	internet_NN1 sales_NN account_VV0 for_IF about_RG 10_MC per_NNU21 cent_NNU22 of_IO all_DB UK_NP1 retail_JJ	2.892	75	%	0	%	0	%	0	%	0	%	os tagged.txt	32%
333	the_AT region_NN1 accounted_VVN for_IF about_RG 20_MC per_NNU21 cent_NNU22 of_IO the_AT software_NN1	3.434	21	%	0	%	0	%	0	%	0	%	os tagged.txt	43%
334	., which_DDQ account_VV0 for_IF about_RG three_MC quarters_NN2 of_IO the_AT retailers_NN2 annual_JJ	5.838	97	%	0	%	0	%	0	%	0	%	os tagged.txt	31%

concordance collocates plot patterns clusters filenames source text notes

9.023 Set tion_NN1 rests_VVZ largely_RR on_II the_AT efforts_NN2 of_IO one_MC1 or_CC two_MC companies_NN2 ., such_II21 as_II22 Infosys_NP1 ., which_DDQ are_VBR impecc



DATA

8

DATA

- Business English (heterogeneous)
- Includes a variety of genres:
 - *Talking/writing about business*
 - Business newspapers, books, radio & tv
 - Academic Business English
 - *Doing business*
 - Written: business reports, emails, press releases...
 - Spoken: job interviews, meetings, telephone calls
- In this study: only language used to talk about business (Nelson 2000)

DATA

- Two 500,000 word corpora (Centre for English Corpus Linguistics, Université catholique de Louvain)
 - **VESPA_Business: Academic Business English**
 - Journal of Management Studies, Marketing Science, Journal of International Business Studies, Journal of Business Venturing, Entrepreneurship, Theory and Practice, Leadership Quarterly
 - **BENews Corpus: Business English News**
 - Financial Times, The Economist, The Independent (business section), Times (Economics)

IDENTIFYING APPROXIMATORS

- ▶ **Approximator:** any device used around numbers to approximate a quantity
- ▶ ‘Any device used around numbers’:
 - (parts of) words, punctuation
 - Added before, after and/or to a number
 - *More or less 20, 50 or something, elevenish,*
- ▶ ‘To approximate a quantity’
 - = to denote imprecision of quantity
- ▶ Were discarded: numbers that did not denote a quantity (year, page, scale, ...)

APPROXIMATORS IN VESPA_B AND BENEWS: RESEARCH QUESTIONS

- Corpus-driven approach vs existing lists?
- Impact of genre?: Less imprecision/approximation in Academic Business English than in Business News ?
 - ‘It is often assumed that the goal of academic writing is clarity and lack of ambiguity’
(Myers 1996)



CORPUS-DRIVEN APPROACH AND EXISTING LISTS OF APPROXIMATORS

13

APPROXIMATORS UNCOVERED BY THE CORPUS-DRIVEN APPROACH: DIVERSITY OF APPROXIMATORS

- 47 types

- Words

- Single words:

- *nearly* \$1 billion (VESPA_B)

- *around* 1,800 staff (BENews)

- revenues *approaching* €12bn (BENews)

- Combinations of words:

- *at most* 2 per cent of economic output (BENews)

- *between 3 and 5 years ago* (VESPA_B)

- a figure *in the region of* 8 per cent (BENews)

APPROXIMATORS UNCOVERED BY THE CORPUS-DRIVEN APPROACH: DIVERSITY OF APPROXIMATORS

○ Parts of words

- Inflectional suffix:

- *thousands of households* (VESPA_B)

- *hundreds of billions of pounds* (BENews)

- Prefix

- *the **multi-** billion euro scramble* (BENews)

○ ≠ words: punctuation

- *a total of EUR 20 billion-25 billion* (BENews)

- *that's going to take two, three, four years* (BENews)

APPROXIMATORS UNCOVERED BY THE CORPUS-DRIVEN APPROACH: DIVERSITY OF APPROXIMATORS

○ Combinations of approximators:

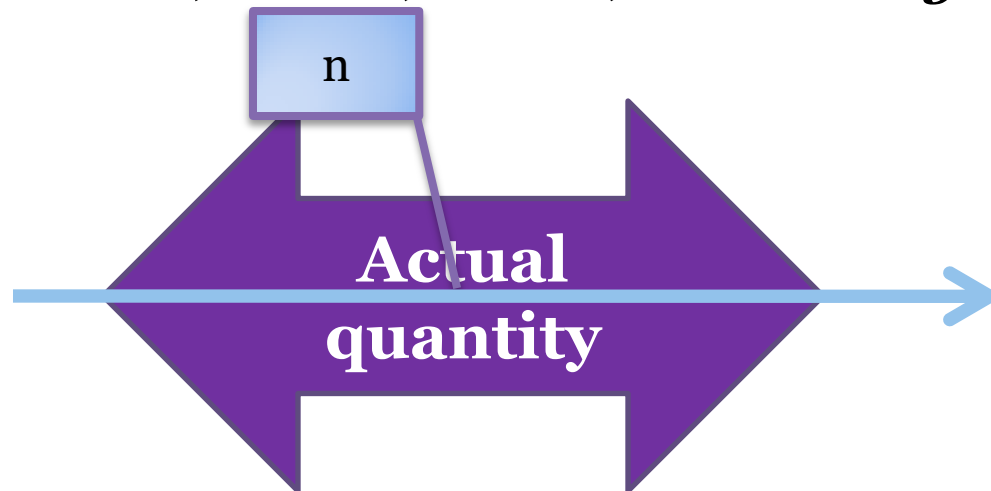
- *there were **about** 5 to 10 top global firms considering entering the Indian market (BENews)*
- *The company never lends **more than** 50 to 80 per cent of the value of the item (BENews)*
- *the available data indicate that over the last two decades this factor has netted the US **at most** \$300bn-\$400bn (BENews)*
- *the goodness-of-fit index (GFI), adjusted goodness-of-fit index (AGFI), normed fit index (NFI), non-normed fit index (NNFI), and comparative fit index (CFI) should be **near** or higher than .90 (VESPA_B)*

GRAMMATICAL DIVERSITY

1. Adverbs (14 types)
2. Prepositions (7 types)
3. Comparative adjective + than (6 types)
4. MORE/LESS + than (5 types)
5. Verbs (4 types)
6. Coordinating conjunction + adverb (3 types)
7. Determiners (2 types)
8. Coordinating conjunction (1 type)
9. Prefix (1 type)
10. Suffix (1 type)
11. Prepositional phrase (1 type)

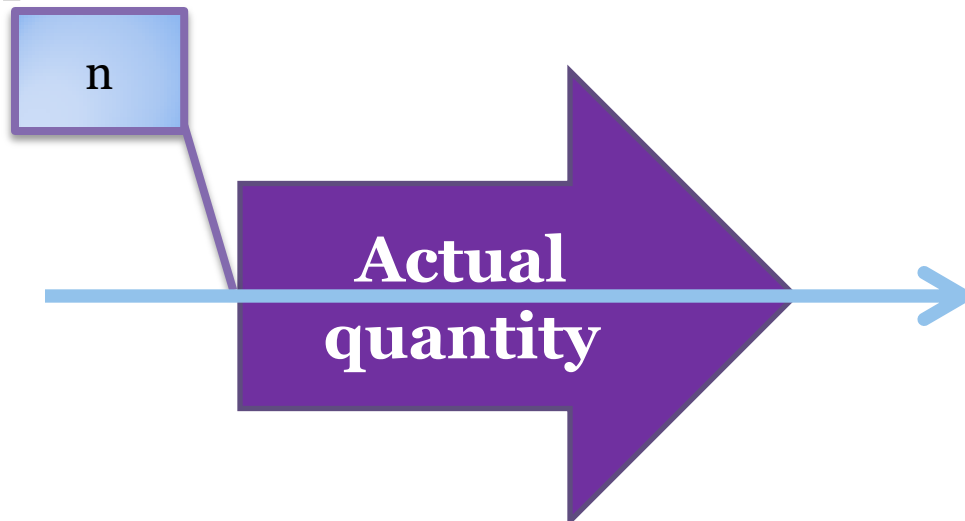
SEMANTIC DIVERSITY

- Lists of approximators in other studies are mostly limited to *about, around, like, approximately, roughly, round, or so*:
 - Actual quantity is larger or smaller than number (expressing approximate amounts)
 - ▶ **9/47** types in VESPA_B and BENews
 - *About, around, approximately, roughly, TO APPROXIMATE, some, circa, in the region of, or so*



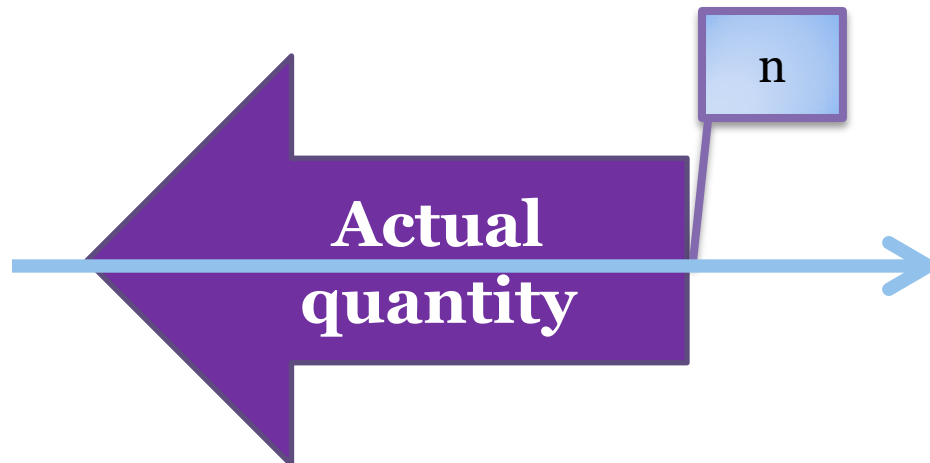
SEMANTIC DIVERSITY

- ▶ Four additional semantic categories were identified in this study
- ▶ Actual quantity $\geq n$ (Expressing minimum amounts): **15/47** types in VESPA_B and BENews
 - ▶ *at least, over, above, beyond, more than, greater than, higher than, longer than, larger than, no less than, TO EXCEED, plural number nouns, several, multi- , or more*



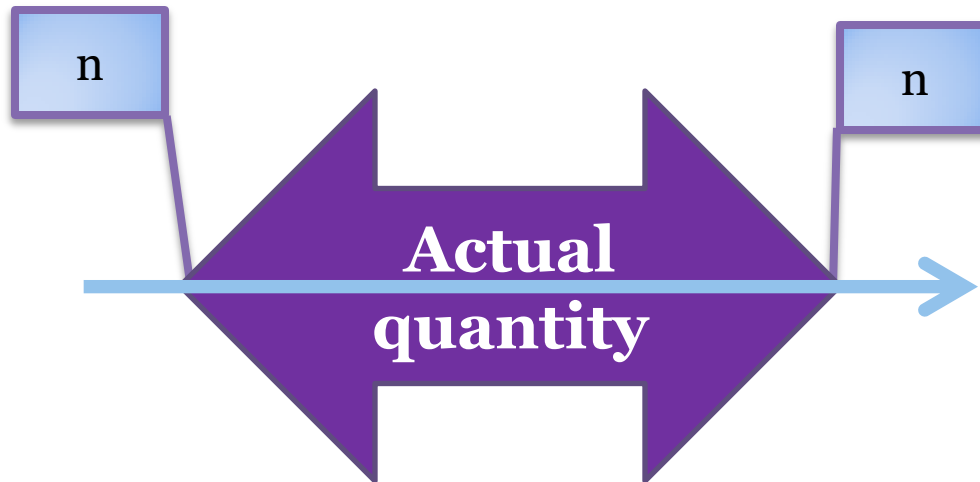
SEMANTIC DIVERSITY

- ▶ Actual quantity $\leq n$ (Expressing maximum amounts): **17/47** types in VESPA_B and BENews
 - *under, below, up to, nearly, near, TO NEAR, close to, almost, towards, TO APPROACH, less than, fewer than, not more than, smaller than, lower than, at most*
 - *Or less*



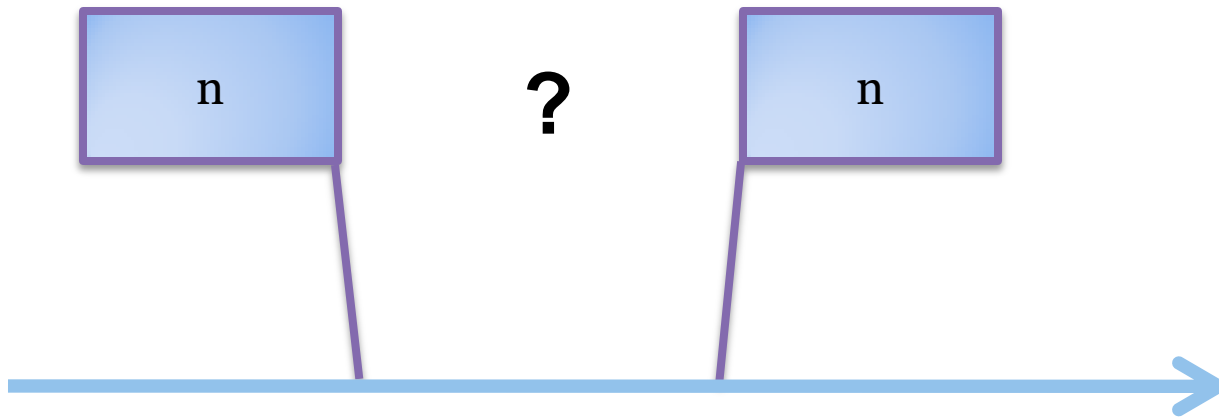
SEMANTIC DIVERSITY

- Actual quantity is between **n** and **n** (interval):
5/47 types in VESPA_B and BENews
 - Between n and n; n to n; n-n; from n to n; n, n, n.*



SEMANTIC DIVERSITY

- ▶ Actual quantity is either **n** or **n: 1/47**
type in VESPA_B and BENews
 - *N or n (often one or two)*



COMPARING THE 47 TYPES IN OUR CORPORA WITH EXISTING LISTS OF APPROXIMATORS


	Number of approximators in existing lists	Number of approximators in existing lists also found in our corpora	Additional types of approximators identified in our corpora (not in existing lists)
Collins Cobuild English Grammar	29	21	26
Cambridge Grammar of English	13	8	39
Channell (1994)	15	14	33
Ruzaitè (2007)	5	4	43

COMPARISON WITH OTHER LISTS OF APPROXIMATORS

- Not found in our corpora:
 - Frequency? Level of formality? Genre-specific?
 - *round about*
 - *(a) minimum/maximum (of)*
 - *or thereabouts*
 - *or something*
 - *-odd*
 - *-ish*
 - *something like*


AROUND APPROXIMATORS: PREFERRED COMPANY

- *Rio's stockmarket value **hovers** around \$35 billion (BENews)*
- *IMRG **estimates** internet sales account for about 10 per cent of all UK retail spending (BENews)*
- *The HBOS investment is undisclosed but is **thought to be** about £2m. (BENews)*
- *The researchers found that a sizable amount of variance (around 32% **on average**) could be accounted by how much a subordinate liked (VEPSA_B)*
- Percentages: *more than* and *about*




COMPARISON OF ACADEMIC BUSINESS ENGLISH (VESPA_B) AND BUSINESS ENGLISH NEWS (BENEWS)

OVERALL PICTURE IN VESPA_BUSINESS AND BENews

	VESPA_B		BENews
Types	35		40
Tokens	388	<	 2042
Ty/to ratio	9,02	>	1,96

PROPORTION OF NUMBERS DENOTING APPROXIMATED QUANTITIES IN VESPA_B AND BENews

	VESPA_B		BENews
Number of 'numbers'	4921	<	14386
Proportion of approximations (tokens)/Numbers	388/4921	<	2042/14386
% of approximations/Numbers	7,88	<	 14,19

APPROXIMATORS IN VESPA_B AND BENews

VESPA_B only (7 types)	Shared (28)		BENews only (12)
<p>greater than</p> <p>lower than</p> <p>not more than</p> <p>TO APPROXIMATE</p> <p>From one to two</p> <p>no less than</p> <p>smaller than</p>	<p>more than</p> <p>about</p> <p>plural number nouns</p> <p>up to</p> <p>around</p> <p>over</p> <p>at least</p> <p>almost</p> <p>nearly</p> <p>between 75% and 80%</p> <p>below</p> <p>less than</p>	<p>approximately</p> <p>above</p> <p>140m-145m</p> <p>one or two</p> <p>close to</p> <p>24,000 to 50,000</p> <p>fewer than</p> <p>TO EXCEED</p> <p>roughly</p> <p>near</p> <p>higher than</p> <p>TO APPROACH</p> <p>30% or more</p> <p>5 or less</p> <p>some</p> <p>longer than</p>	<p>10% or so</p> <p>under</p> <p>several</p> <p>multi-</p> <p>towards</p> <p>beyond</p> <p>TO NEAR</p> <p>at most</p> <p>circa</p> <p>larger than</p> <p>two, three, four</p> <p>in the region of</p>

47 types found in VESPA_B + BENews

TOP TEN APPROXIMATORS IN VESPA_B AND BENEWS

Top ten approximators in VESPA_B	Number of tokens in VESPA_B (total: 388 tokens)	Top ten approximators in BENEWS	Number of tokens in BENEWS (total: 2042 tokens)
at least	53	more than	447
more than	44	about	325
over	41	plural number nouns (dozens, hundreds, thousands, millions, billions)	234
about	37	up to	173
n or more	29	around	135
approximately	27	almost	89
greater than	16	over	89
less than	16	nearly	85
almost	15	at least	63
between n and n	14	some	58

APPROXIMATORS IN VESPA_B AND BENews

- Genre differences => different approximators?
 - *Approximately*: ‘Generally confined to rather specific official, semi-official, and scientific registers’ (Channell 1994)
 - 27 in **VESPA_B** > 7 in **BENews**
 - *Two, three, four* (**BENews** only): reported speech!
 - Business News Reporting: hybrid genre => also speech!
 - Plural number nouns: more in **BENews** + combined
 - 234 in **BENews** > 14 in **VESPA_B**
 - Amounts of money

The slide features a dark blue background with a decorative left side consisting of vertical stripes in various shades of blue and purple. Several circles of different sizes and colors (purple, blue, and white) are scattered on the left side, some overlapping the stripes.

CONCLUSIONS

32

CONCLUSIONS

- Corpus-driven approach: diversity
 - Number of types (> existing lists)
 - Grammar
 - Semantics
- More approximators in **BENews** than VESPA_B (proportion of numbers approximated and tokens)
- Genre-specific approximators

REFERENCES

- ▶ Carter, R. & McCarthy, M. (2006) *Cambridge Grammar of English. A comprehensive guide: Spoken and Written English: Grammar and Usage*. Cambridge: Cambridge University Press.
- ▶ Channell, J. (1994) *Vague Language*. Oxford: Oxford University Press
- ▶ Collins Cobuild English Grammar (1998). London: Harper Collins.
- ▶ Jucker, A.H., Smith, S.W. & Lüdge, T. (2003) Interactive aspects of vagueness in conversation. *Journal of Pragmatics* 35:1737-1769.
- ▶ Koester, A. (2007) 'About twelve thousand or so': Vagueness in North American and UK offices. In Cutting, J., *Vague Language Explored*. New-York: Palgrave MacMillan.
- ▶ Myers, G. (1996) Strategic vagueness in academic writing. In: Ventola, E. & A. Mauranen (eds.) *Academic Writing: Intercultural and Textual Issues*. Amsterdam: John Benjamins.
- ▶ Nelson, M. (2000) *A corpus-based study of the Lexis of Business English and Business English teaching materials*. Unpublished PhD thesis. University of Manchester. Manchester.
- ▶ Ruzaitè, J. (2007) *Vague Language in Educational Settings - Quantifiers and Approximators in British and American English*. Frankfurt am Main: Lang.
- ▶ Tognini-Bonelli, E. (2001) *Corpus Linguistics at Work: Studies in Corpus Linguistics*. Amsterdam: John Benjamins.