Large-scale evaluation of dependency-based DSMs: Are they worth the effort? Lapesa & Evert, EACL 2017 – extended supplementary material.

1	Dist	ribution of Performance	2
	1.1	Multiple choice: TOEFL	2
	1.2	Prediction of similarity ratings: RG65	3
	1.3	Prediction of similarity ratings: WS353	3
	1.4	Clustering: Almuhareb-Poesio	4
	1.5	Clustering: BATTIG	4
	1.6	Clustering: MITCHELL	5
	1.7	Clustering: ESSLLI	5
	1.8	Semantic priming, multiple choice: SYN	6
	1.9	Semantic priming, multiple choice: ANT	6
	1.10	Semantic priming, multiple choice: COH	7
	1.11	Semantic priming, multiple choice: FPA	7
	1.12	Semantic priming, multiple choice: BPA	8
	1.13	Semantic priming, multiple choice: GEK	8
_			•
2	Expl	anatory Power of DSM Parameters	9
	2.1	TOEFL	9
	2.2	Ratings	10
	2.3		11
	2.4	Semantic priming, paradigmatic relations (SYN, ANT, COH)	12
	2.5	Semantic priming, syntagmatic relations (FPA, BPA, GEK)	13
3	Effe	ct plots	14
3	Effe 3.1	ct plots TOEFL	14 14
3	Effe 3.1 3.2	ct plots TOEFL	14 14 21
3	Effe 3.1 3.2 3.3	ct plots TOEFL WS353 RG65	14 14 21 28
3	Effe 3.1 3.2 3.3 3.4	ct plots TOEFL WS353 RG65 AP	14 14 21 28 35
3	Effe 3.1 3.2 3.3 3.4 3.5	ct plots TOEFL WS353 RG65 AP BATTIG	14 14 21 28 35 43
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL	14 14 21 28 35 43 52
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI	14 14 21 28 35 43 52 60
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN	14 14 21 28 35 43 52 60 69
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT	14 14 21 28 35 43 52 60 69 76
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH	 14 14 21 28 35 43 52 60 69 76 83
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA	14 14 21 28 35 43 52 60 69 76 83 91
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA<	14 14 21 28 35 43 52 60 69 76 83 91 98
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA GEK	14 14 21 28 35 43 52 60 69 76 83 91 98 105
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA GEK	14 14 21 28 35 43 52 60 69 76 83 91 98 105
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13 Best	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA GEK TOEFI	14 14 21 28 35 43 52 60 69 76 83 91 98 105 112
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13 Best 4.1	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA GEK TOEFL Potiere	14 14 21 28 35 43 52 60 69 76 83 91 98 105 112 112
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13 Best 4.1 4.2	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA GEK TOEFL Ratings Clustering	14 14 21 28 35 43 52 60 69 76 83 91 98 105 112 112 113
3	Effe 3.1 3.2 3.3 3.4 3.5 3.6 3.7 3.8 3.9 3.10 3.11 3.12 3.13 Best 4.1 4.2 4.3 4.4	ct plots TOEFL WS353 RG65 AP BATTIG MITCHELL ESSLLI SYN ANT COH FPA BPA GEK models TOEFL Ratings Clustering Citering	14 14 21 28 35 43 52 60 69 76 83 91 98 105 112 112 113 114

1 Distribution of Performance

1.1 Multiple choice: TOEFL



Figure 1.3: Unreduced, typed



Figure 1.4: Reduced, typed

1.2 Prediction of similarity ratings: RG65



Figure 1.7: Unreduced, typed

1.3 Prediction of similarity ratings: WS353



Figure 1.9: Unreduced, filtered



Figure 1.11: Unreduced, typed



Figure 1.8: Reduced, typed



Figure 1.10: Reduced, filtered



Figure 1.12: Reduced, typed

1.4 Clustering: Almuhareb-Poesio



Figure 1.15: Unreduced, typed



Figure 1.14: Reduced, filtered Min: 0.13; Max: 0.78; Mean: 0.53; Stidev :0.116



Figure 1.16: Reduced, typed

1.5 Clustering: BATTIG



Figure 1.17: Unreduced, filtered



Figure 1.19: Unreduced, typed



Figure 1.18: Reduced, filtered



Figure 1.20: Reduced, typed

1.6 Clustering: MITCHELL





Figure 1.23: Unreduced, typed



Figure 1.22: Reduced, filtered Min: 0.25; Max: 0.95; Mean: 0.64; Stdev :0.117



Figure 1.24: Reduced, typed

1.7 Clustering: ESSLLI



Figure 1.25: Unreduced, filtered



Figure 1.27: Unreduced, typed



Figure 1.26: Reduced, filtered



Figure 1.28: Reduced, typed

1.8 Semantic priming, multiple choice: SYN





Figure 1.31: Unreduced, typed

1.9 Semantic priming, multiple choice: ANT



Figure 1.33: Unreduced, filtered



Figure 1.35: Unreduced, typed



Figure 1.30: Reduced, filtered



Figure 1.32: Reduced, typed



Figure 1.34: Reduced, filtered



Figure 1.36: Reduced, typed

1.10 Semantic priming, multiple choice: COH



Figure 1.39: Unreduced, typed

1.11 Semantic priming, multiple choice: FPA



Figure 1.41: Unreduced, filtered



Figure 1.43: Unreduced, typed



Figure 1.38: Reduced, filtered



Figure 1.40: Reduced, typed



Figure 1.42: Reduced, filtered



Figure 1.44: Reduced, typed

1.12 Semantic priming, multiple choice: BPA





Figure 1.47: Unreduced, typed

1.13 Semantic priming, multiple choice: GEK



Figure 1.49: Unreduced, filtered



Figure 1.51: Unreduced, typed



Figure 1.46: Reduced, filtered



Figure 1.48: Reduced, typed



Figure 1.50: Reduced, filtered



Figure 1.52: Reduced, typed

2 Explanatory Power of DSM Parameters

2.1

TOEFL



Figure 2.53: Unreduced, filtered



Figure 2.55: Unreduced, typed



Figure 2.54: Reduced, filtered



Figure 2.56: Reduced, typed

2.2 Ratings



Figure 2.57: Unreduced, filtered



Figure 2.58: Reduced, filtered



Figure 2.59: Unreduced, typed



Figure 2.60: Reduced, typed

2.3 Clustering



Figure 2.61: Unreduced, filtered



Figure 2.62: Reduced, filtered



Figure 2.63: Unreduced, typed



Figure 2.64: Reduced, typed

Semantic priming, paradigmatic relations (SYN, ANT, COH) 2.4



Figure 2.65: Unreduced, filtered

score

metric

corpus

parser

dep.type

dep.style

path.length

context.dim

transformation



Figure 2.66: Reduced, filtered



Figure 2.67: Unreduced, typed



Figure 2.68: Reduced, typed

2.5 Semantic priming, syntagmatic relations (FPA, BPA, GEK)



Figure 2.69: Unreduced, filtered



Figure 2.70: Reduced, filtered



Figure 2.71: Unreduced, typed



Figure 2.72: Reduced, typed

3 Effect plots

3.1 TOEFL

3.1.1 Main Effects



3.1.2 Interactions: filtered, unreduced







Figure 3.14: TOEFL, filtered, unreduced











45-5000 10000 20000 50000 100000 Figure 3.18: TOEFL, filtered, unreduced

70-

65-

60-

55-

50-





Figure 3.19: TOEFL, filtered, unreduced

70

65

60-

55

50

45

bnc



ukwac

wacky Figure 3.21: TOEFL, filtered, unreduced



3.1.3 Interactions: filtered, reduced

Figure 3.23: TOEFL, filtered, reduced

















3.1.4 Interactions: typed, unreduced





Figure 3.31: TOEFL,typed,unreduced





Figure 3.34: TOEFL,typed,unreduced





Figure 3.35: TOEFL,typed,unreduced





Figure 3.37: TOEFL,typed,unreduced















3.2 WS353

3.2.1 Main Effects









Figure 3.14: WS353, filtered, unreduced











































3.2.4 Interactions: typed, unreduced





Figure 3.31: WS353,typed,unreduced











3.2.5 Interactions: typed, reduced













Figure 3.41: WS353,typed,reduced

Figure 3.42: WS353,typed,reduced



3.3 RG65

3.3.1 Main Effects









Figure 3.14: RG65, filtered, unreduced









Figure 3.18: RG65, filtered, unreduced



3.3.3 Interactions: filtered, reduced















Figure 3.25: RG65, filtered, reduced















Figure 3.30: RG65,typed,unreduced













Figure 3.37: RG65,typed,unreduced



















3.4 AP

3.4.1 Main Effects



3.4.2 Interactions: filtered, unreduced













Figure 3.16: AP, filtered, unreduced



36




3.4.3 Interactions: filtered, reduced











37





Figure 3.26: AP, filtered, reduced





Figure 3.28: AP,filtered,reduced







Figure 3.31: AP, filtered, reduced







Figure 3.33: AP,typed,unreduced











Figure 3.37: AP,typed,unreduced



Figure 3.38: AP,typed,unreduced

3.4.5 Interactions: typed, reduced













41









Figure 3.46: AP,typed,reduced



Figure 3.47: AP,typed,reduced

3.5 BATTIG

3.5.1 Main Effects



3.5.2 Interactions: filtered, unreduced





















Figure 3.20: BATTIG, filtered, unreduced



Figure 3.21: BATTIG, filtered, unreduced









Figure 3.23: BATTIG, filtered, reduced





















Figure 3.31: BATTIG, filtered, reduced











3.5.4 Interactions: typed, unreduced





Figure 3.36: BATTIG,typed,unreduced









Figure 3.40: BATTIG,typed,unreduced



Figure 3.41: BATTIG,typed,unreduced



Figure 3.42: BATTIG,typed,unreduced







Figure 3.44: BATTIG,typed,reduced





0.9

0.8

0.7

0.6





Figure 3.48: BATTIG,typed,reduced



0.9

0.8-

0.7

0.6

٨٨



Figure 3.51: BATTIG,typed,reduced







50

100

ò



3.6 MITCHELL

3.6.1 Main Effects



3.6.2 Interactions: filtered, unreduced





Figure 3.14: MITCHELL, filtered, unreduced











Figure 3.18: MITCHELL, filtered, unreduced

3.6.3 Interactions: filtered, reduced

















Figure 3.24: MITCHELL, filtered, reduced













Figure 3.30: MITCHELL, filtered, reduced

Number of Skipped Dimensions * Transformation

Figure 3.28: MITCHELL, filtered, reduced

Number of Skipped Dimensions * Metric

55



Figure 3.31: MITCHELL, filtered, reduced

3.6.4 Interactions: typed, unreduced







Figure 3.33: MITCHELL,typed,unreduced





Figure 3.36: MITCHELL,typed,unreduced





Figure 3.37: MITCHELL,typed,unreduced



3.6.5 Interactions: typed, reduced



































3.7 ESSLLI

3.7.1 Main Effects



3.7.2 Interactions: filtered, unreduced

















Figure 3.18: ESSLLI, filtered, unreduced







Figure 3.20: ESSLLI, filtered, unreduced



Figure 3.21: ESSLLI, filtered, unreduced







Figure 3.23: ESSLLI, filtered, reduced









Dependency group * Metric









Figure 3.27: ESSLLI, filtered, reduced



0.75





Figure 3.31: ESSLLI, filtered, reduced







Figure 3.33: ESSLLI, filtered, reduced



















Figure 3.41: ESSLLI,typed,unreduced

















Figure 3.46: ESSLLI,typed,unreduced



















Figure 3.52: ESSLLI,typed,reduced













Figure 3.56: ESSLLI,typed,reduced



Figure 3.57: ESSLLI,typed,reduced

3.8 SYN

3.8.1 Main Effects











Figure 3.17: SYN, filtered, unreduced



Figure 3.14: SYN, filtered, unreduced





Figure 3.18: SYN, filtered, unreduced













Figure 3.24: SYN, filtered, reduced









Figure 3.26: SYN, filtered, reduced



Figure 3.28: SYN, filtered, reduced

3.8.4 Interactions: typed, unreduced





Figure 3.30: SYN,typed,unreduced










3.8.5 Interactions: typed, reduced





Figure 3.37: SYN,typed,reduced









Figure 3.41: SYN,typed,reduced

.





Figure 3.44: SYN,typed,reduced

3.9 ANT

3.9.1 Main Effects





















3.9.3 Interactions: filtered, reduced







Score * Transformation





Figure 3.26: ANT, filtered, reduced

100

ò 50





Score * Metric









Figure 3.28: ANT, filtered, reduced



Figure 3.30: ANT, filtered, reduced

3.9.4 Interactions: typed, unreduced



Figure 3.31: ANT,typed,unreduced



Figure 3.32: ANT,typed,unreduced





Figure 3.34: ANT,typed,unreduced



Figure 3.35: ANT,typed,unreduced



3.9.5 Interactions: typed, reduced





Figure 3.39: ANT,typed,reduced







Figure 3.41: ANT,typed,reduced





3.10 COH

3.10.1 Main Effects









Figure 3.14: COH, filtered, unreduced











Figure 3.18: COH, filtered, unreduced





3.10.3 Interactions: filtered, reduced



Figure 3.21: COH, filtered, reduced















Figure 3.26: COH, filtered, reduced



Figure 3.27: COH, filtered, reduced













Figure 3.33: COH, filtered, reduced







Figure 3.35: COH,typed,unreduced











Figure 3.39: COH,typed,unreduced









Figure 3.41: COH,typed,unreduced



Figure 3.43: COH,typed,unreduced

3.10.5 Interactions: typed, reduced





Figure 3.45: COH,typed,reduced





Figure 3.47: COH,typed,reduced









3.11 FPA

3.11.1 Main Effects

















Figure 3.18: FPA, filtered, unreduced

















3.11.3 Interactions: filtered, reduced





Number of Skipped Dimensions * Transformation

transformation

90-









Figure 3.28: FPA, filtered, reduced





3.11.4 Interactions: typed, unreduced





Figure 3.32: FPA,typed,unreduced











Figure 3.36: FPA,typed,unreduced





































3.12 BPA

3.12.1 Main Effects





Figure 3.13: BPA, filtered, unreduced







Figure 3.15: BPA, filtered, unreduced







Figure 3.18: BPA, filtered, unreduced

3.12.2 Interactions: filtered, unreduced

























Figure 3.26: BPA, filtered, reduced





Figure 3.27: BPA, filtered, reduced

Figure 3.28: BPA, filtered, reduced



3.12.4 Interactions: typed, unreduced























Figure 3.37: BPA,typed,unreduced













3.12.5 Interactions: typed, reduced





Figure 3.43: BPA,typed,reduced

Number of Skipped Dimensions * Metric







90

85

Figure 3.46: BPA,typed,reduced

3.13 GEK

3.13.1 Main Effects



3.13.2 Interactions: filtered, unreduced





Figure 3.14: GEK, filtered, unreduced





3.13.3 Interactions: filtered, reduced











Figure 3.18: GEK, filtered, reduced





Figure 3.22: GEK, filtered, reduced







Number of Skipped Dimensions * Transformation transformatio none 🔶 log 📥 root 🔶 sigmo 80-70 60-50 ò 100

Figure 3.24: GEK, filtered, reduced



Figure 3.26: GEK, filtered, reduced

3.13.4 Interactions: typed, unreduced



Figure 3.27: GEK,typed,unreduced






Figure 3.30: GEK,typed,unreduced



3.13.5 Interactions: typed, reduced









Figure 3.36: GEK,typed,reduced



Figure 3.37: GEK,typed,reduced











4 Best models

4.1 TOEFL

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	wacky	stanford	core	ccproc	1	50k	z-score	none	cosine	rank	85.00
typed	wacky	stanford	core	basic	2	100k	z-score	none	cosine	rank	83.75
typed	wacky	stanford	core	basic	2	50k	z-score	none	cosine	rank	83.75

Table 4.1: TOEFL, unreduced, best models - Filtered vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	ukwac	malt	external	basic	2	5k	t-score	log	500	100	cosine	rank	93.75
filtered	ukwac	stanford	external	basic	3	5k	t-score	log	500	900	cosine	rank	93.75
filtered	wacky	stanford	external	basic	2	10k	simple-ll	log	500	900	cosine	rank	93.75
typed	ukwac	stanford	external	basic	4	50k	MI	none	100	700	cosine	dist	91.25
typed	ukwac	stanford	external	basic	4	50k	MI	none	100	900	cosine	dist	91.25
typed	ukwac	stanford	external	basic	1	100k	MI	root	100	900	cosine	dist	91.25

Table 4.2: TOEFL, reduced. Filtered (3 runs tied for best result) vs. Typed (3 runs tied for best result)

4.2 Ratings

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	r
filtered	wacky	malt	external	ccproc	1	50k	MI	none	cosine	rank	0.88
typed	wacky	malt	core	ccproc	1	100k	z-score	none	manhattan	rank	0.80

Table 4.3: RG65, unreduced, best models - Filtered vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	r
filtered	ukwac	malt	core	basic	4	50k	MI	none	50	500	cosine	rank	0.88
typed	wacky	malt	core	basic	1	100k	z-score	log	100	900	cosine	rank	0.87

Table 4.4: RG65, reduced, best models - Filtered vs. Type

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	r
filtered	ukwac	stanford	external	ccproc	4	50k	z-score	none	cosine	rank	0.71
typed	ukwac	stanford	external	basic	1	100k	z-score	root	cosine	rank	0.59

Table 4.5:	WS353,	unreduced,	best models	- Filtered	vs.	Typed
------------	--------	------------	-------------	------------	-----	-------

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	r
filtered	ukwac	stanford	core	ccproc	3	100k	z-score	root	50	900	cosine	rank	0.72
typed	ukwac	stanford	external	basic	1	100k	MI	none	50	900	cosine	rank	0.66

Table 4.6: WS353, reduced, best models - Filtered vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	purity
filtered	ukwac	malt	external	basic	1	100k	z-score	log	cosine	rank	0.75
typed	wacky	stanford	external	ccproc	1	100k	z-score	none	manhattan	rank	0.75
typed	wacky	malt	external	ccproc	1	100k	z-score	root	cosine	rank	0.75
typed	wacky	stanford	external	ccproc	1	100k	z-score	none	manhattan	rank	0.75

Table 4.7: AP, unreduced, best models - Filtered vs. Typed (3 runs tied for best result)

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	purity
filtered	wacky	malt	core	ccproc	1	20k	t-score	none	0	900	manhattan	rank	0.75
typed	ukwac	stanford	external	basic	1	100k	z-score	root	0	300	cosine	rank	0.78

Table 4.8: AP, reduced, best models - Filtered vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	purity
filtered	bnc	malt	external	basic	2	100k	z-score	none	manhattan	rank	0.98
filtered	ukwac	malt	core	basic	2	100k	simple-ll	log	cosine	rank	0.98
filtered	wacky	stanford	external	basic	2	50k	z-score	none	manhattan	dist	0.98
typed	ukwac	stanford	core	ccproc	1	100k	Dice	root	cosine	rank	0.95

Table 4.9: BATTIG, unreduced, best models - Filtered (46 runs tied for best result, 3 hand-picked examples shown) vs.Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	purity
filtered	bnc	malt	core	basic	4	50k	z-score	root	0	500	cosine	rank	0.99
filtered	ukwac	malt	core	ccproc	4	100k	z-score	none	100	500	manhattan	rank	0.99
filtered	ukwac	malt	external	basic	1	100k	freq	log	50	300	cosine	dist	0.99
typed	ukwac	stanford	core	ccproc	1	100k	z-score	root	50	100	cosine	rank	1.00

Table 4.10: BATTIG, reduced, best models - Filtered (520 runs tied for best result, 3 hand-picked examples shown) vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	purity
filtered	bnc	stanford	external	basic	1	50k	MI	none	cosine	rank	0.91
filtered	wacky	stanford	external	basic	1	100k	simple-ll	log	manhattan	rank	0.91
filtered	ukwac	stanford	external	basic	1	50k	z-score	log	manhattan	rank	0.91
filtered	wacky	stanford	external	ccproc	1	100k	z-score	none	manhattan	dist	0.91
typed	bnc	malt	external	basic	1	20k	t-score	sigmoid	cosine	rank	0.89
typed	bnc	malt	external	basic	1	50k	MI	root	cosine	rank	0.89

Table 4.11: ESSLLI, unreduced, best models - Filtered (4 runs tied for best result) vs. Typed (2 runs tied for best result)

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	purity
filtered	ukwac	stanford	core	basic	1	100k	simple-ll	log	50	700	cosine	dist	0.98
filtered	ukwac	stanford	core	basic	1	50k	tf.idf	root	50	500	cosine	dist	0.98
filtered	wacky	malt	external	basic	3	100k	z-score	none	0	700	manhattan	rank	0.98
typed	ukwac	stanford	external	basic	1	100k	simple-ll	log	50	100	cosine	rank	0.98

Table 4.12: ESSLLI, reduced, best models - Filtered (29 runs tied for best result 3 hand-picked examples shown) vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	purity
filtered	bnc	external	malt	basic	1	100k	simple-ll	log	cosine	rank	0.93
filtered	ukwac	external	malt	basic	3	10k	simple-ll	root	cosine	rank	0.93
filtered	bnc	external	stanford	basic	2	50k	simple-ll	root	cosine	rank	0.93
typed	bnc	external	stanford	basic	1	100k	z-score	none	manhattan	dist	0.90
typed	bnc	external	stanford	basic	1	50k	z-score	none	manhattan	dist	0.90
typed	bnc	external	stanford	basic	2	100k	z-score	none	cosine	rank	0.90

Table 4.13: MITCHELL, unreduced, best models - Filtered (33 runs tied for best result, 3 hand-picked examples shown) vs. Typed (3 runs tied for best result).

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	purity
filtered	bnc	stanford	external	basic	2	20k	z-score	root	0	700	cosine	rank	0.97
typed	bnc	malt	external	ccproc	1	100k	Dice	root	50	100	cosine	rank	0.95
typed	bnc	stanford	external	basic	1	100k	z-score	root	50	300	cosine	rank	0.95

Table 4.14: MITCHELL, reduced, best models - Filtered vs. Typed (2 runs tied for best result)

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	wacky	stanford	external	ccproc	1	100k	z-score	log	cosine	rank	97.94
filtered	wacky	stanford	external	ccproc	1	50k	simple-ll	log	cosine	rank	97.94
filtered	wacky	stanford	external	ccproc	1	50k	z-score	root	cosine	rank	97.94
filtered	wacky	stanford	external	ccproc	1	5k	z-score	none	manhattan	rank	97.94
typed	wacky	stanford	external	ccproc	1	100k	z-score	none	cosine	rank	96.56
typed	wacky	stanford	external	basic	2	100k	z-score	none	cosine	rank	96.56

Table 4.15: SYN, unreduced, best models - Filtered (4 runs tied for best result) vs. Typed (2 runs tied for best result)

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	ukwac	stanford	external	ccproc	1	50k	z-score	log	100	900	cosine	rank	99.31
typed	ukwac	stanford	external	basic	1	100k	tf.idf	log	50	900	cosine	rank	97.25

Table 4.16: SYN, reduced, best models - Filtered vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	bnc	malt	core	basic	4	10k	frequency	none	manhattan	rank	100.00
filtered	ukwac	stanford	external	basic	3	100k	z-score	none	cosine	rank	100.00
filtered	wacky	stanford	core	basic	3	20k	t-score	none	cosine	rank	100.00
typed	bnc	malt	core	ccproc	1	50k	MI	none	cosine	rank	100.00
typed	ukwac	malt	external	ccproc	1	10k	Dice	none	cosine	rank	100.00
typed	wacky	malt	core	basic	4	50k	simple-ll	log	manhattan	rank	100.00

Table 4.17: ANT, unreduced, best models - Filtered (5387 runs tied for best result, 3 hand-picked examples shown) vs. Typed (1469 runs tied for best result, 3 hand-picked examples shown).

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	bnc	malt	core	ccproc	4	50k	MI	none	0	700	cosine	rank	100.00
filtered	wacky	stanford	external	ccproc	2	100k	MI	log	0	900	cosine	rank	100.00
filtered	ukwac	stanford	external	ccproc	1	5k	MI	root	0	300	cosine	dist	100.00
typed	ukwac	malt	core	basic	4	20k	tf.idf	log	100	700	cosine	dist	100.00
typed	wacky	stanford	external	ccproc	3	5k	simple-ll	sigmoid	50	500	manhattan	rank	100.00
typed	wacky	stanford	external	ccproc	1	5k	t-score	cosine	50	900	cosine	rank	100.00

Table 4.18: ANT, reduced, best models - Filtered (23209 runs tied for best result, 3 hand-picked examples shown) vs. Typed: (805 runs tied for best result, 3 hand-picked examples shown).

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	wacky	stanford	external	ccproc	1	50k	simple-ll	root	cosine	rank	100.00
filtered	ukwac	stanford	external	basic	1	50k	Dice	none	cosine	rank	100.00
filtered	bnc	stanford	core	basic	1	100k	simple-ll	log	cosine	rank	100.00
typed	wacky	stanford	external	ccproc	1	10k	simple-ll	none	manhattan	dist	100.00
typed	wacky	stanford	external	ccproc	2	50k	z-score	none	cosine	rank	100.00
typed	ukwac	stanford	core	ccproc	1	100k	MI	sigmoid	cosine	rank	100.00

Table 4.19: COH, unreduced, best models - Filtered (1139 runs tied for best result, 3 hand-picked examples shown) vs. Typed (721 runs tied for best result, 3 hand-picked examples shown).

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	ukwac	stanford	external	basic	2	100k	MI	log	50	500	cosine	rank	100.00
filtered	wacky	stanford	external	ccproc	2	5k	z-score	log	50	900	cosine	rank	100.00
filtered	ukwac	malt	core	basic	1	10k	MI	none	50	700	manhattan	rank	100.00
typed	bnc	stanford	core	ccproc	1	10k	MI	root	0	900	cosine	rank	100.00
typed	wacky	stanford	external	ccproc	3	50k	z-score	root	0	700	cosine	rank	100.00
typed	ukwac	stanford	external	ccproc	1	100k	t-score	sigmoid	0	700	cosine	rank	100.00

Table 4.20: COH, reduced, best models - Filtered (8237 runs tied for best result, 3 hand-picked examples shown) vs. Typed (2617 runs tied for best result, 3 hand-picked examples shown).

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	ukwac	stanford	core	basic	2	100k	Dice	none	cosine	rank	97.22
filtered	ukwac	stanford	core	basic	2	50k	z-score	none	cosine	rank	97.22
filtered	wacky	malt	external	basic	4	50k	z-score	none	cosine	rank	97.22
typed	wacky	stanford	core	basic	1	50k	MI	none	cosine	rank	88.19
typed	wacky	stanford	external	basic	1	100k	MI	none	cosine	rank	88.19
typed	wacky	stanford	external	basic	1	50k	z-score	none	manhattan	rank	88.19

Table 4.21: FPA, unreduced, best models - Filtered (9 runs tied for best result, 3 hand-picked examples shown). Typed (5 runs tied for best result, 3 hand-picked examples shown)

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	wacky	malt	external	ccproc	3	20k	z-score	none	0	900	cosine	dist	97.22
filtered	ukwac	stanford	core	ccproc	4	10k	tf.idf	none	50	50	manhattan	rank	97.22
filtered	wacky	stanford	external	basic	3	10k	z-score	none 0	900	cosine	rank		97.22
typed	ukwac	stanford	external	basic	1	100k	z-score	root	50	900	cosine	rank	90.28

Table 4.22: FPA, reduced, best models -. Filtered (15 runs tied for best result, 3 hand-picked examples shown) vs. Typed.

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	ukwac	malt	core	basic	3	100k	z-score	root	cosine	rank	97.75
filtered	ukwac	malt	external	basic	2	100k	Dice	log	cosine	rank	97.75
filtered	ukwac	malt	external	ccprocessed	4	20k	z-score	none	cosine	rank	97.75
typed	ukwac	malt	external	basic	1	100k	z-score	root	cosine	rank	92.13

Table 4.23: BPA, unreduced, best models - Filtered (43 runs tied for best result, 3 hand-picked examples shown) vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	bnc	malt	external	ccprocessed	3	100k	MI	none	50	700	cosine	rank	98.88
filtered	ukwac	malt	core	basic	4	100k	simple-ll	log	50	700	cosine	rank	98.88
filtered	ukwac	malt	core	ccprocessed	3	100k	Dice	root	50	500	cosine	rank	98.88
typed	ukwac	stanford	core	basic	2	100k	Dice	root	100	900	cosine	rank	95.51
typed	ukwac	stanford	core	basic	2	100k	Dice	root	100	900	cosine	rank	95.51

Table 4.24: BPA, reduced, best models - Filtered (365 runs tied for best result) vs. Typed (2 models tied for best result)

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	metric	rel.ind	accuracy
filtered	ukwac	stanford	external	basic	4	50k	z-score	none	cosine	rank	95.54
typed	ukwac	stanford	external	basic	1	100k	z-score	none	cosine	rank	87.13

Table 4.25: GEK, unreduced, best models - Filtered vs. Typed

	corpus	parser	d.group	d.style	p.len	c.dim	score	transf	d.skip	n.dim	metric	rel.ind	accuracy
filtered	ukwac	malt	external	basic	4	10k	MI	none	50	700	manhattan	rank	95.79
filtered	ukwac	malt	external	basic	4	50k	Dice	root	50	900	cosine	dist	95.79
filtered	ukwac	malt	external	basic	4	50k	Dice	root	50	900	cosine	rank	95.79
filtered	ukwac	stanford	external	ccproc	3	50k	Dice	root	50	900	cosine	dist	95.79
typed	ukwac	malt	core	basic	2	100k	tf.idf	log	50	300	cosine	dist	89.60
typed	ukwac	malt	external	basic	2	100k	frequency	log	50	900	cosine	dist	89.60

Table 4.26: GEK, reduced, best models - Filtered (4 runs tied for best result). Typed (2 runs tied for best result).