Supplementary Materials: Stimulus Additional Information

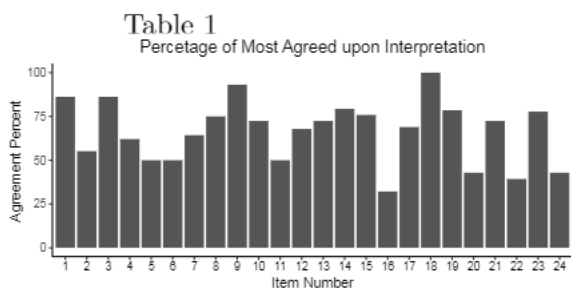Anonymous

Anonymous

Experiment 1 Stimuli

The number of characters and syllables in the content words of the NC and PP was controlled (see Table 1). The head nouns (i.e., last word in NC and first word in PP) all had frequencies between 20 and 400 per million words in both the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). This range enabled using words that were complex enough to mirror scientific texts.

We conducted two online surveys to assess the suitability of individual items for the study. The first survey tested the comprehensibility of 31 NCs and their counterpart PPs. Items were divided across two lists so that no participant saw a given item in both its NC and PP form. Forty monolingual native English speakers from the USA rated the comprehensibility of each item on a Likert scale (1-7). Seven items were removed from the initial 31 on the basis of having the highest or lowest mean ratings, or a particularly high mean rating difference between the NC and PP equivalents. This ensured that the NC and PP variants of a sentence did not differ in difficulty, and also that neither the NC nor the PP were too difficult or too easy within each length condition.

The second survey tested whether the resulting 24 NCs were ambiguous without prior context, given that a primary assumption of the study was that they were somewhat ambiguous. Fifty-seven native English speakers from the USA selected an interpretation for each of the 24 NCs from a set of six possible answers, as illustrated in (1).

(1) "College graduate career earnings lifetime potential" is:
    ( ) the lifetime potential of career earnings from college graduates
    ( ) the potential earnings over the lifetime of college graduate careers
    ( ) the potential for lifetime earnings in careers of college graduates
    ( ) this phrase cannot be interpreted because it is ambiguous
    ( ) this phrase cannot be interpreted because it is meaningless
    ( ) other: _____

The first three choices were always potential interpretations of the NC in which the component NPs were transposed, but only one was the correct interpretation (here, the first option). Only one item evoked 100% agreement for the correct interpretation, while the other items ranged between 32% and 86% agreement (see Figure 1). We concluded that the NCs were not unambiguously interpretable by the participants, and thus were suitable items for our study.

Table 1

Percetage of Most Agreed upon Interpretation

Agreement Percent

100
75
50
25
0

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24

Item Number

...the NC and PP in Experiment 1.

	Length 4		Length 6	
es	No. Characters	No. Syllables	No. Characters	
	4-13	1-4	4-12	
	28-38	13-15	41-48	
	35-42	16-19	51-60	

Figure 1. Percentage of participants who selected the most-agreed-upon interpretation for each of the critical items.

Experiment 2 Stimuli

Similarly to Experiment 1, the number of characters and syllables in the content words of the NC and PP was controlled (see Table 2). Items were also controlled for collocations by three native speakers.

The nouns used for the stimuli cannot be used as verbs. In order to find head nouns that fulfill this condition, some of them have a low frequency in the Contemperary Corpus of American English (COCA) or on the British National Corpus (BNC). Frequencies of all words on the BNC and on the COCA are presented in Tables 3 and 4, respectively.

The PPs were constructed as a "translation" of the NCs into the PP structure. This "translation" required a choice for the linking preposition between the PP nouns. For each item of length 4 this required the choice of 3 prepositions. No choice was needed for items

Table 2

Range of characters and syllables in the NC and PP in Experiment 2.

	Head Noun	NCs		PPs	
		Length 3	Length 4	Length 3	Length 4
No. Characters	5-12	26-30	34-39	34-46	48-62
No. Syllables	1-5	7-12	10-15	10-16	14-21

Table 3

Frequency of nouns used in Experiment 2 on the British National Corpus.

Position 1		Position 2		Position 3		Position 4		Item average
Word	Frequency	Word	Frequency	Word	Frequency	Word	Frequency	
currency	33.29	inflation	43.16	constraint	6.39	action	214.76	74.40
factory	43.87	automation	3.80	legislation	68.29	advice	102.06	54.51
storage	28.99	capacity	57.20	maximum	50.09	indicator	8.18	36.12
employee	30.66	insurance	69.24	payment	53.20	reduction	47.38	50.12
nutrition	5.03	fact	362.95	disclosure	10.08	regulation	25.31	100.84
office	243.97	technology	115.78	adjustment	13.95	period	239.12	153.21
alcohol	29.72	producer	17.50	advertisement	10.52	group	406.32	116.02
border	37.83	taxation	24.79	consequence	33.77	summary	27.98	31.09
transit	8.62	energy	119.55	performance	127.31	standard	124.81	95.07
product	108.75	quality	159.84	assurance	17.97	department	173.50	115.02
highway	9.20	construction	61.60	equipment	87.16	subsidy	7.18	41.29
utilities	4.76	monopoly	16.11	operation	98.27	condition	81.66	50.20
internet	0.87	commerce	15.63	growth	126.80	expectation	12.80	39.03
company	393.26	acquisition	26.07	oversight	1.63	council	305.90	181.72
capital	131.91	allocation	18.21	efficiency	36.31	analysis	130.33	79.19
welfare	47.31	fraud	17.23	investigation	49.75	committee	185.93	75.06
aluminum	10.13	shipment	3.02	application	98.01	paperwork	4.75	28.98
military	109.24	aircraft	60.72	transaction	22.48	proposal	41.23	58.42
gender	19.47	employment	105.62	equality	15.13	movement	130.52	67.69
investment	106.84	portfolio	15.01	diversity	13.83	strategy	60.26	48.99
business	347.15	revenue	40.07	maximization	1.28	principle	80.21	117.18
candidate	37.60	campaign	91.70	speech	75.92	presentation	31.19	59.10
potato	8.51	shortage	14.52	problem	282.51	management	215.84	130.35
healthcare	2.52	policy	254.43	disapproval	4.90	response	117.06	94.73
history	188.46	education	256.83	modernization	2.82	agenda	23.00	117.78
machinery	23.92	rental	4.61	agreement	127.54	negotiation	11.46	41.88
minority	33.51	voter	2.79	participation	26.31	statistic	31.62	23.56
hurricane	4.75	relief	63.22	organization	82.48	donation	5.98	39.11
politics	72.39	newspaper	47.23	coverage	21.51	commentary	8.46	37.40
region	96.97	independence	42.62	referendum	13.04	outcome	36.30	47.23
Average	73.98		71.04		52.64		96.37	73.51
Min	0.87		2.79		1.28		4.75	23.56
Max	393.26		362.95		282.51		406.32	181.72

with length 3, since each length-3 item is a shorter form of a length-4 item and used the exact same prepositions. In total, since there were 28 items, this led to $3 \times 28 = 84$ preposition choices. The preposition "of" was always the preferred preposition (used in 61 out of the 84 cases), but the prepositions "for" (used in 16 cases), "on" (3 cases), "to" (2 cases) "in" (1 case) and "at" (1 case) were also used when "of" did not sound natural. In most cases, the preposition was followed by the determiner "the" and the noun in its singular form; but in a few cases the noun was used in the plural (7 cases), or the determiner "the" was dropped (23 cases).

We conducted two online surveys to assess the suitability of the individual items for the study. The first survey tested the comprehensibility of the 28 NCs and their counterpart PPs, with lengths 3 and 4. Items were divided across four lists so that no participant saw a given item in more than one condition (3-NC, 4-NC, 3-PP or 4-PP). Thirty-eight L1 native English speakers from the USA rated the comprehensibility of each item on a Likert scale (1-7). In order to compare the items in the different conditions, we calculated the average comprehensibility of NC version and subtracted from it the average comprehensibility of the PP version. Figure 2 shows the comprehensibility difference for each item and each length. No item had a difference higher than 1.8, and the average difference was 0.2.

The second survey tested whether the resulting 28 NCs were ambiguous without prior context, given that a primary assumption of the study was that they were somewhat

Table 4

Frequency of nouns used in Experiment 2 on the Contemporary Corpus of American English.

Position 1 Word	Frequency	Position 2 Word	Frequency	Position 3 Word	Frequency	Position 4 Word	Frequency	Item average
currency	16.46	inflation	22.64	constraint	2.57	action	144.87	46.64
factory	27.42	automation	2.34	legislation	46.11	advice	57.01	33.22
storage	28.53	capacity	47.56	maximum	26.80	indicator	8.78	27.92
employee	28.84	insurance	86.38	payment	18.86	reduction	30.60	41.17
nutrition	18.12	fact	354.81	disclosure	10.06	regulation	24.10	101.77
office	254.05	technology	153.73	adjustment	16.94	period	133.80	139.63
alcohol	41.63	producer	26.68	advertisement	3.68	group	387.79	114.95
border	57.50	taxation	3.97	consequence	15.33	summary	19.69	24.12
transit	13.24	energy	161.20	performance	139.43	standard	95.76	102.41
product	77.08	quality	115.98	assurance	6.26	department	192.09	97.85
highway	38.76	construction	65.26	equipment	65.97	subsidy	4.23	43.56
				operation	57.52	condition	65.54	35.48
				growth	128.10	expectation	10.62	61.77
				oversight	10.60	council	94.03	108.57
				efficiency	24.53	analysis	125.71	61.13
				investigation	71.74	committee	114.38	62.31
				application	35.73	paperwork	7.69	15.45
				transaction	7.97	proposal	37.12	75.91
				equality	14.09	movement	112.27	53.90
investment	70.57	portfolio	14.70	diversity	35.75	strategy	77.58	49.65
business	326.83	revenue	36.11	maximization	0.37	principle	33.90	99.30
candidate	60.26	campaign	151.35	speech	86.28	presentation	22.25	80.04
potato	15.85	shortage	10.83	problem	300.37	management	112.37	109.86
healthcare	14.26	policy	200.64	disapproval	3.46	response	121.71	85.02
history	280.83	education	317.88	modernization	4.97	agenda	36.55	160.06
machinery	8.86	rental	10.98	agreement	75.83	negotiation	8.87	26.14
minority	43.14	voter	11.64	participation	43.96	statistic	33.18	32.98
hurricane	22.26	relief	47.57	organization	99.87	donation	6.36	44.02
politics	105.90	newspaper	56.47	coverage	51.03	commentary	10.61	56.00
region	101.25	independence	36.61	referendum	7.49	outcome	35.12	45.12
Average	83.12		69.13		47.06		72.15	67.86
Min	8.86		2.34		0.37		4.23	15.45
Max	326.83		354.81		300.37		387.79	160.06

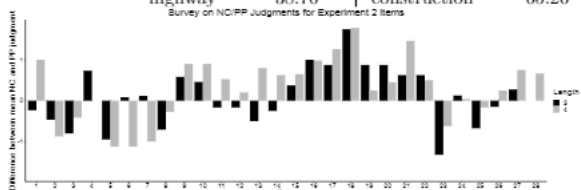Survey on NC/PP Judgments for Experiment 2 items

Figure 2. Difference in comprehensiblility between items in their PP and NC version.
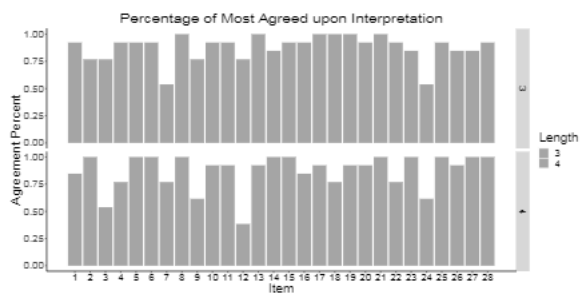
Figure 3. Percentage of participants who selected the most-agreed-upon interpretation for each of the critical items..

ambiguous. The NCs of lengths 3 and 4 were divided into two lists, so that each list contained either items with length 3 or 4. Twenty-seven native English speakers from the USA selected an interpretation for each of the 28 NCs from a set of seven possible answers, as illustrated in (2). The interpretations were constructed so that their head noun was different in each choice. For items with length 3, the fourth interpretation correctly identified the head noun, but had an incorrect ordering of the other nouns (compare the first and the fourth choice below).

(2) "inflation constraint action" means:
    ( ) the action for the constraint of inflation
    ( ) the constraint for the action of inflation
    ( ) the inflation for the constraint of action
    ( ) the action for the inflation of constraint
    ( ) this phrase cannot be interpreted because it is ambiguous
    ( ) this phrase cannot be interpreted because it is meaningless
    ( ) other _____

Figure 3 shows the frequency with which participants chose the most common interpretation for items with length 3 and 4. In most cases the most-agreed upon interpretation was very dominant.