Apparent Hurford constraint obviations are based on scalar implicatures: An argument based on frequency counts[***].

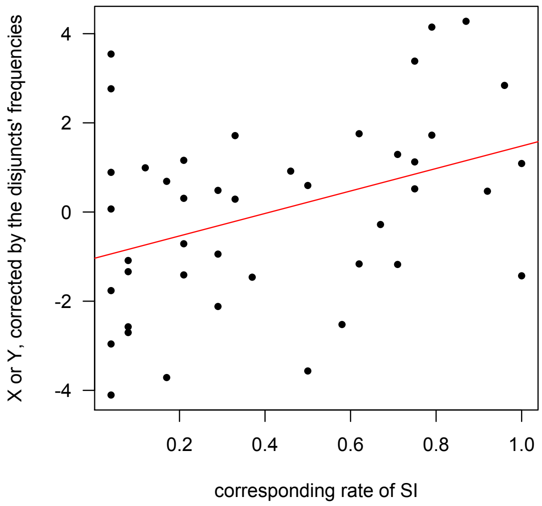
Emmanuel Chemla, CNRS

According to Hurford (1974), a disjunction is not felicitous if one of the disjuncts entails the other, as in (1). There are however systematic obviations to this constraint (Gazdar 1979). Effectively, it has been proposed that a sentence such as (2) is felicitous, despite the fact that *X*="reading some books" is entailed by *Y*="reading all books", because *X* can be strengthened and interpreted as *X and not-Y*="reading some but not all books", by means of a local scalar implicature (see discussion in Chierchia, Fox & Spector 2013).

(1) John is in France or in Paris.
(2) John read some or all of the books. (2') X or Y
(3) John read some of the books. → John did not read all the books. (3') X → not-Y

Using a plain inferential task, van Tiel et al. (2013) have gathered quantitative data revealing important variability in the derivation rate of inferences such as (3)/(3') for different *<X,Y>* scales. If the obviation analysis in terms of scalar implicature is correct, and if van Tiel et al.'s data indeed reflects (even to a small extent) the derivation rate of scalar implicatures, we expect that the felicity or the frequency of (2') should co-vary with the derivation rate of the inference in (3') as given by van Tiel et al.

For each scale *<X,Y>* investigated in van Tiel et al., we collected a (noisy) estimate of the frequencies of *X*, of *Y* and of the disjunction *X or Y*, as the number of hits obtained from a google search of these elements (between quotation marks). We assume that, despite the noise, the frequency of *X or Y*, corrected by the frequencies of each of the disjuncts, approximates its felicity, i.e. the potential for the pair *(X,Y)* to escape from Hurford's constraint in a disjunction.



The raw data are in Table 1 . Technically, we first ran a linear model by which the log-frequency of the disjunction *X or Y* is predicted by the log-frequencies of both disjuncts. We used the residuals obtained from this model as a corrected frequency of the disjunction. Crucially, we ran a second model to see whether the rate of derivation of the inference (3'), as reported in van Tiel et al. (2013) accounts for some of the remaining variability in this corrected frequency of the disjunction.

We obtain an overall significant correlation: $r^2 = .15$, $F(1,41)=7.0$, $p=.012$. (One may prefer to apply a Poisson regression to model the counts of disjuncts based on the log-frequencies of each disjunct to extract the residuals in the first step described above, such an analysis also yields a significant correlation in the second step: $r^2 = .091$, $F(1,41)=4.1$, $p=.0497$). This correlation shows that the more participants are willing to derive an inference *X→not-Y* (as measured by van Tiel et al.'s task), the more the corresponding *X or Y* disjunction occurs. We interpret this result as further evidence that (i) Hurford's constraint is active and favors disjunctions in which the first disjunct does not entail the other and (ii) that apparent obviations of this constraint are reinforcements of *X* into *X and not-Y,* which occur within a sentence (the first disjunct) in a comparable way as they occur at the sentential level.

Overall, the conjunction of old generalizations on disjunctions and recent discussions about the status of scalar implicatures in the grammar predicted the covariation between two sets of rather different data:

[***]. Thanks to Danny Fox, Raj Singh, Philippe Schlenker and Benjamin Spector for discussion. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.313610 and was supported by ANR-10-IDEX-0001-02 and ANR-10-LABX-0087.

frequency counts and inferential preferences. A correlation between these data emerged out of noisy estimates for both sides of the equation (van Tiel et al's data may represent scalar implicatures to a small extent and google estimates of frequency are rough). The emergence of the correlation despite this noisy environment further validates the original theoretical motivations for looking after this correlation.

Chierchia, G., Fox, D. & Spector, B. (2013). The Grammatical View of Scalar Implicatures and the Relationship between Semantics and Pragmatics. In P. Portner, C. Maienborn, and K. von Heusinger (eds), Handbook of Semantics. Mouton de Gruyter, New York, NY.

Gazdar, G. (1979). Pragmatics: Implicature, Presupposition and Logical Form. Academic Press, New York, NY.

Hurford, J. R. (1974). "Exclusive or Inclusive Disjunction " *Foundation of Language*,11: 409-411.

van Tiel, B., van Miltenburg, E., Zevakhina, N. & Bart Geurts, B. (2013). Scalar diversity. Ms. University of Nijmegen.

X	**Y**	**Rate of Sis X->not-Y**	**Freq of X**	**Freq of Y**	**Freq of "X or Y"**	**Corrected frequency (residuals)**
cheap	free	100	1.04E+09	1.15E+10	15800000	1.08799974
sometimes	always	100	7.15E+08	2.23E+09	432000	-1.430650938
some	all	96	4.93E+09	1.97E+10	316000000	2.839654934
possible	certain	92	1.62E+09	7.78E+08	2770000	0.466740092
may	have to	87	8.17E+09	1.59E+09	494000000	4.27751033
difficult	impossible	79	5.21E+08	3.16E+08	34100000	4.14705019
rare	extinct	79	5.59E+08	2.10E+07	774000	1.723482463
few	none	75	1.87E+09	1.04E+09	3710000	0.519550365
may	will	75	8.17E+09	8.84E+09	492000000	3.384092471
warm	hot	75	5.49E+08	3.92E+09	6320000	1.123629612
hard	unsolvable	71	2.25E+09	1190000	22800	-1.175870815
low	depleted	71	2.19E+09	16700000	1040000	1.291729998
allowed	obligatory	67	5.90E+08	16900000	96400	-0.280416128
scarce	unavailable	62	33500000	2.22E+08	475000	1.756379281
try	succeed	62	3.19E+09	1.17E+08	309000	-1.164134644
palatable	delicious	58	6180000	7.04E+08	4200	-2.52305388
like	love	50	9.38E+09	5.16E+09	389000	-3.565035249
memorable	unforgettable	50	94900000	58500000	142000	0.595198636
good	excellent	46	5.24E+09	1.38E+09	12100000	0.916762672
good	perfect	37	5.24E+09	1.65E+09	1230000	-1.462064357
cool	cold	33	1.70E+09	7.00E+08	2260000	0.288164511
hungry	starving	33	2.07E+08	26800000	470000	1.713643712
adequate	good	29	1.21E+08	5.24E+09	1520000	0.485036982
dislike	loathe	29	1.18E+09	6210000	45400	-0.943806456
unsettling	horrific	29	7110000	22800000	1160	-2.118404212
believe	know	21	1.03E+09	4.01E+09	749000	-1.410672597
participate	win	21	3.06E+08	1.23E+09	2500000	1.1591853
start	finish	21	3.69E+09	6.01E+08	3440000	0.307209967
wary	scared	21	25400000	1.33E+08	26000	-0.711767262
big	enormous	17	3.91E+09	6.26E+08	65400	-3.712498385
old	ancient	17	3.87E+09	3.30E+08	3800000	0.688022425
snug	tight	12	27100000	4.45E+08	278000	0.99152558
attractive	stunning	8	2.79E+08	3.29E+08	28500	-2.574097035
intelligent	brilliant	8	2.03E+08	2.55E+08	90700	-1.08735116
pretty	beautiful	8	1.22E+09	1.82E+09	594000	-1.337887629
special	unique	8	2.93E+09	1.23E+09	213000	-2.703125635
content	happy	4	4.80E+09	2.01E+09	5960000	0.068012295
dark	black	4	1.27E+09	4.71E+09	197000	-2.959377675
funny	hilarious	4	1.03E+09	1.13E+08	7960	-4.104692313
silly	ridiculous	4	1.47E+08	87700000	2010000	2.764260944
small	tiny	4	4.81E+09	4.59E+08	445000	-1.762418829
tired	exhausted	4	2.50E+08	44500000	4280000	3.542814075
ugly	hideous	4	1.76E+08	12200000	124000	0.889668623

Table 1: Raw data reporting the two members of each scale <*X,Y*>, the corresponding derivation rate of scalar implicature X→not-Y (from van Tiel et al. 2013), a rough estimate of the frequencies of X, of Y and of "X or Y" (as the number of hits obtained from a google search of these expressions), and the corrected frequency of the disjunction (as the residuals of a regression of the log-frequency of the disjunction by the log-frequencies of X and of Y).