METHODS

A synthetic description of several methodological details is available in Table 1 of the main paper.

EXPERIMENT 1

Participants. Thirty-nine students of the University of Padua (26 female, mean age 22.43, standard deviation (SD) = 3.44, range = 18-30) took part in the experiment for course credit. Participants in this and in Experiments 2-to-8 were Italian native speakers, had normal or corrected to normal vision and took part only in one experiment.

Materials. Cue-word selection was based on a pre-test questionnaire. Twenty participants from the same population that did not take part in the main experiment were asked to judge each of 70 words as referring to objects usually appearing in the top or bottom portion of the visual field using a 7-point scale (1 = lowest; 7 = highest). Reliability was high, intra-class correlation coefficient (ICC)=.98, with a 95% confidence interval from .97 to .99. The ICC and the confidence intervals were calculated using a two-way random effects model on average measures with an absolute agreement definition using the psych package (Revelle, 2017) with the R program (R Development Core Team, 2016). The two groups of 12 words (i.e., cued Top and Bottom location words) differed for the top/bottom judgments (t (22) = 28.18, p < .001) but were equated for frequency, length, orthographic uniqueness point, orthographic and phonological neighborhood size, mean frequency of the orthographic and phonological neighbors (*p*s > .11, see materials below).

Procedure. The experiment was run at the University of Padova. An experimental trial consisted of the following events. A fixation cross was shown in the centre of the screen for 250 ms, followed by the centrally presented cue-word [1] and then the target stimulus at varying SOAs. The target was centered horizontally but at the top or bottom of the screen (5° from the center) and remained until response. The inter-trial interval was 500 ms. The G and H keys used for

---

[1] Cue-word duration is reported in Table 1 of the paper. As in Experiment 3 (and differently from Experiments 1-2) by Estes et al. (2008) no context words preceded the presentation of the cue-words.

responding and were covered by stickers with a square and a circle on them. These two letters in the Italian keyboard appear next to each other. The correspondence between the key (G or H) and the target (a square or a circle) was counterbalanced between participants. Differently from the Estes et al. paradigm in which participants used the two hands for responding, participants responded with the middle and index finger of their dominant hand. Participants were tested individually in a soundproof, uniformly lit room. Stimulus presentation and response times (RTs) were controlled by E-Prime 2 (Psychology Software Tools, Inc., Sharpsburg, PA) E-prime software has been used to control all subsequent experiments except for Experiment 4. Stimuli were presented on an Intel compatible computer running Windows XP with a Acer cathode ray tube monitor (diagonal 15''). Participants were seated approximately 60 cm from the screen. Stimuli were presented on a display with a resolution of 768x1980 pixels. Prime and target stimuli were presented in white on a black background in Courier New font, size 18.

Design. All stimuli were presented in three blocked conditions of 96 trial each, according to the duration of the interval between the offset of the cue-word and the onset of the stimulus (i.e., 150, 450, and 900 ms). The order of block presentation was counterbalanced between subjects according to a Latin square design. There was a short pause after each block. Trial order within each block was random. Cue-word type (top or bottom), target position (top or bottom) and target itself (a circle or a square) were fully crossed and balanced, so that each cue-word was repeated 4 times in each block, followed by a circle or a square presented either at the top or bottom position. Four new cue-words without a clear spatial cue location were presented in sixteen practice trials before the main experiment.

EXPERIMENT 2 AND 3

Participants. Forty-one participants (36 female, mean age 20.49, SD = 6.49, range = 18-29) took part in Experiment 2 and 20 (19 female, mean age 19.5, SD = 0.98, range = 19-23) took part in Experiment 3 for course credit. They were all students at the University of Padova.

Materials, Procedure and Design. The stimuli and procedure were identical to those of Experiment 1, except that number words (i.e., three, four, five, six, seven, eight, nine) were added as catch trials. Catch trials were presented randomly among the experimental trials with a total of 6 catch trials per block. On catch trials, participants had to respond both to the number word, by reading it aloud, and to following target stimulus, as in all other trials. In Experiment 3, the letters J and K were used as target stimuli and as response keys. These two letters in the Italian keyboard appear next to each other in the same row, and are not easily associated to semantic categories since they are found extremely rarely in the Italian language.

EXPERIMENT 4

Participants. Eighteen students of the University of Trento (10 female, mean age 21.8, SD = 2.47, range = 19-27) took part in the experiment for course credit.

Materials, Procedure and Design. The experiment was run at the University of Trento. Sixty-four new Italian words were selected as cue-words. The two sets of cue-words were matched on several psycholinguistic variables (ps > .21, see materials below). The fixation cross was presented for 500 ms, followed by the cue-word, and then the target letter, which appeared 7° above or below the centre of the screen. Participants responded with the index fingers of the two hands, as in Estes et al. (2008). The 128 experimental trials were presented in a single block and trial order was randomized. The full crossing of cue-word type, type of target and location of the target was obtained using two lists, so that each word was associated with a target letter in a position for half of the participants and in the other position for the other half. For instance, half of the participants saw the cue-word *head* with the target M located at the top and with the target C located at the

bottom; the other half of the participants saw the same cue-word with the reverse combination: M located at the bottom and C located at the top. Each participant was randomly assigned to one of the two lists, so that each cue word was presented two times to each participant. Eight new cue-words without a clear spatial cue location were presented in eight practice trials. Stimuli were presented on a computer Dell OptiPlex 520, Pentium 4 3.0 GHz., with Dell cathode ray tube monitor, diagonal 21''. Stimuli were presented on a display with a resolution of 768x1980 pixels. Prime and target stimuli were presented in black on a white background in Arial font, size 18. Stimulus presentation and response recording was controlled by Experiment Builder software (SR Research).

EXPERIMENT 5 AND 6

Participants. Twenty students of the University of Padua (19 female, mean age 19.4, SD= 0.9, range = 18-22) took part in Experiment 5 and 24 (21 female, mean age 19.5, SD= 0.7, range = 18-21) took part in Experiment 6 for course credit.

Materials. For selecting the cue-word, 10 participants from the same population but that did not take part in the main experiment were asked to judge each of 160 Italian words as referring to the top or bottom portion of the visual field using a 7-point scale (1 = lowest; 7 = highest). Fifty-six Italian words which corresponded to 28 words ranked 'highest' and 28 words ranked 'lowest' were selected as cue-words. Reliability was high, intra-class correlation coefficient (ICC)=.95, with a 95% confidence interval from .93 to .97. The two groups of cue-words differed in the top/bottom judgments (t (54) = 37.93, p < .001) but not in in several psycholinguistic variables (*p*s > .23, see materials below). In addition, filler trials in which the cue-word had no spatial connotation (e.g., *paper*) were randomly intermingled with experimental trials, as in Estes et al. (2008, Experiment 3). Filler trials were intermixed with experimental trials and in both types of trials participants were not required to perform any task on the cue-words.

Procedure and Design. The experiment was run at the University of Padova. Target stimuli were the letters X or O. They appeared at 5° and 8° (as in Estes et al., 2008) from the center of the screen in Experiment 5 and 6, respectively. The response keys were "X" and "O" on the keyboard in Experiment 5 and "C" and "M" in Experiment 6. As in Experiment 4 and as in Estes et al. (2008), participants responded with the index finger of the two hands. In Experiment 6, participants pressed the space bar to start each trial (i.e. self paced) in order to further approach the procedure used in Estes et al. (2008). Each cue word was followed by one of the two targets located in the top or bottom position. As in Experiment 4, in order to prevent word repetition, two lists were constructed so that each word was associated with a target letter in a position for half of the participants and in the other position for the other half. In this manner each word was presented 2 times for a total of 224 trials, resulting from 112 words (56 experimental cue words and 56 filler words) repeated twice. For all other details we used the same procedure as in Experiment 1.

EXPERIMENT 7 and 8

Participants. Twenty-five students of the University of Padua (23 female, mean age 19.2, SD= 0.89, range = 18-23) took part in Experiment 7 and 40 in Experiment 8 (34 female, mean age 22.6, SD= 4.52, range = 18-34). The participants of Experiment 8 were randomly assigned to the two instruction conditions.

Materials. These were the same as in Experiments 5 and 6. However, the instructions given to participants were manipulated at the beginning of the experimental session. In Experiment 7 and in the biased condition of Experiment 8 the instructions were: "At the centre of the screen you will see a cross. Fix your eyes on that point. Afterwards a word will appear. Some of the words refer to objects that usually occur in the top part of the visual field, and others refer to objects that usually occur in the bottom part of the visual field. After each word, a letter X or a letter O will appear in the top or in the bottom portion of the display. Wherever the stimuli appear and

whatever word precedes them, the task is to simply press the X key if you see the letter X and the O key if you see the letter O. Press the space bar to start the next trial. Try to be as quick and accurate as possible in your responses." Neutral instruction used in the not-biased condition of Experiment 8 (and in Experiments 1 to 6) were: "At the centre of the display you will see a cross. Fix your eyes on that point. Afterwards a word will appear, and after that, a letter. The task is to press the X key if you see the letter X and the O key if you see the letter O. Press the space bar to start the next trial. Try to be as quick and accurate as possible in your responses".

Procedure and Design. Both experiments were run at the University of Padova using the same procedure as in Experiment 6. Note that the biased condition of Experiment 8 was an exact replication of Experiment 7. The neutral condition of Experiment 8 was an exact replication of Experiment 6.

EXPERIMENT 9

Participants. Forty students from Trent University, Canada (28 female, mean age 24.68, SD= 7.62, range = 19-59) took part in the Experiment for course credit. They were English native speakers with normal or corrected to normal vision.

Materials . The cue-words and the filler stimuli were the same as used by Estes and coll. (2008).

Procedure and Design. The experiment was run at Trent University, Canada. Stimuli were presented on a Dell Vostro 420 computer. Stimuli were presented on a Dell E207WFP display with a resolution of 1680 x 1050 pixels. The keys used for the response were the X and the O of the keyboard. All other details were the same as in Experiment 6.

EXPERIMENT 10

Participants. Twenty students from the University of Padova (18 female, mean age 24.25, SD= 3.04, range = 20-32) took part in the Experiment for course credit.

Materials . The 56 cue-words of experiment 5-8 were used. Two lists were created. In one list, half of the top and half of the bottom words were presented as prime stimuli and the other half as target stimuli. In the other list, words reversed the combination prime/target. In each list, each target (and prime word) was paired with four different words (prime or target): two top and two bottom words.

Procedure. The 128 experimental trials were presented in a single block and trial order was randomized. The same experimental parameters as in Experiments 5-8 were used: Prime word was presented for 100 ms and after a blank of 50 ms the target word was presented until response. Eight new target words (4 bottom and 4 top) and 8 new prime words (without a clear spatial location) were presented in 16 practice trials. Participants were instructed to indicate as fast and as accurately as possible whether the target word denotes a concept that usually appears at the top or at the bottom of the visual field by pressing the C or the M keys. Participants responded with the index fingers of the two hands. The correspondence between the keys and the target type was counterbalanced between participants.

RESULTS

Table 2. The mean and standard deviation (SD) for response times (RT) in milliseconds and percentage error (%E) for the *Location Cue Congruency (LCC) effect* and the *Semantic Priming effect*. For Experiment 1, 2 and 3 the data are reported separately for each block, according to the duration of the interval between the offset of the cue word and the onset of the stimulus (i.e., 150, 450, and 900 ms). For Experiment 8 data are reported separately for the two instruction conditions (i.e., neutral and biased).

LCC EFFECT		Congruent				Incongruent			
		RTs	*SD*	*%E*	*SD-E*	*RTs*	*SD*	*%E*	*SD-E*
Experiment 1									
	Block 150	490	86	2.0	1.3	491	83	2.0	1.7
	Block 450	498	86	2.3	1.9	499	85	2.5	1.9
	Block 900	498	99	2.5	2.2	499	95	2.4	1.3
Experiment 2									
	Block 150	571	87	4.0	3.8	572	82	3.0	4.5
	Block 450	551	96	3.0	3.1	545	92	3.0	4.4
	Block 900	557	87	2.0	3.2	553	86	2.0	3.5
Experiment 3									
	Block 150	532	43	5.2	4.0	532	38	5.0	2.7
	Block 450	525	45	4.2	2.9	533	60	4.0	3.0
	Block 900	519	28	3.9	2.7	522	34	4.2	2.9
Experiment 4		482	97	3.4	5.1	480	103	2.9	4.2
Experiment 5		494	73	6.1	3.8	489	74	5.9	4.4
Experiment 6		523	78	5.7	4.6	521	66	5.5	3.6
Experiment 7		526	61	5.2	3.7	516	58	6.5	4.5
Experiment 8									
	Neutral Ins.	527	182	4.5	4.8	525	176	4.7	4.8
	Biased Ins.	554	173	3	2.8	564	187	4.8	4.5
Experiment 9		684	133	6.2	5.1	684	136	6.1	4.3
SEMANTIC PRIMING									
Experiment 10		732	106	4	5	755	111	5	6.1

Analyses were performed only on experimental trials. The Congruency between the location cued by the word and the location of the target stimulus (congruent vs. incongruent) was analyzed through paired t-test treating participants (t1) and items (t2) as random s factors. In the RTs analysis, the main effect of Congruency was only marginally significant in the participant analysis of Experiment 7 (t1 (24) = 2.06, p = .05; t2 (55) = 1.88, p = .065). In the rest of experiments the main effect of Congruency was not significant (ts < 1) (see Table 2). In the analysis of the error rates, the main effect of Congruency was not significant (ts < 1). Two participants who did not name more than 2 catch trials in Experiment 2 were excluded from the analysis. Furthermore, we

tested LCC and FI effects by comparing a null model without predictors with a model containing the LCC and the FI predictors separately, using the package ''lme4'' (Bates, Maechler, & Bolker, 2011) with the R program (2016). Models were compared using the likelihood ratio test and taking into consideration the Bayesian Information Criterion (BIC; Schwarz, 1978). The last column of Table 1 reports the Bayes Factor's (BF) approximation using the formula exp(ΔBIC/2) (Raftery, 1995), where ΔBIC indicates the difference between the null model and other models containing the LCC and the FI predictors, respectively. The Approx. BF indicates the relative evidence of a model, with higher Approx. BF corresponding to a better fit of the tested model respect to the null one.

For Experiments 1, 2 and 3 ANOVA analyses were run with Block (150, 450 and 900) and Congruency as within-participant factors. The main effect of Block was significant in Experiment 2 ($F1(2, 76) = 3.12$, $p < .05$, $\eta_p^2 = .07$; $F2 (2, 46) = 21.71$, $p < .01$, $\eta_p^2 = .48$) and significant in the item analysis only in Experiment 1 ($F1 < 1$; $F2 (2, 46) = 7.12$, $p < .003$, $\eta_p^2 = .23$) and in Experiment 3 ($F1 (2, 38) = 1.13$, $p = .33$, $\eta_p^2 = .05$; $F2 (2, 46) = 5.88$, $p < .01$, $\eta_p^2 = .2$). The interaction between Block and Congruency was not significant ($Fs < 1$). In the analysis of the error rates, the main effect of Block failed to reach the standard significance threshold ($ps > .09$). The interaction between the factors Congruency and Block was also not significant ($Fs < 1$). As can be seen in Table 2, RTs in Experiment 2 were slower than in Experiment 1 ($F1(1, 38) = 16.28$, $p < .001$, $\eta_p^2 = .3$; $F2 (1, 23) = 2750.36$, $p < .001$, $\eta_p^2 = .99$), probably reflecting the additional task to be accomplished on the number words.

For Experiment 8, ANOVA analyses were run with Congruency as within subject factor and Type of Instructions (Biased vs. Neutral) as between subject factor. The main effect of Type of Instruction was significant in the item analysis only ($F1(1, 38) = 1.19$, $p = .28$, $\eta_p^2 = .03$; $F2 (1, 55) = 61.96$, $p < .01$, $\eta_p^2 = .53$). The interaction between the factors Type of Instructions and Congruency

was not significant ($F1(1, 38) = 3.12$, $p = .08$, $\eta_p^2 = .07$; $F2 (1, 55) = 2.23$, $p = .14$, $\eta_p^2 = .03$). There were no effects in the error rates analysis ($ps > .14$).

In order to assess statistical sensitivity, we computed the power of observing an effect the same size as the one reported by Estes et al. (Experiment 3, 32 ms). Power was computed using the function *power.t.test* of the R software (2016). We obtained a $1-\beta$ value $> .97$ in all experiments which clearly indicated that the probability of a Type II error was very low.

For Experiment 10 analyses were run comparing congruent (i.e. both prime and target indicating a "high" or a "low" concept) vs. incongruent (i.e. prime indicating a "high" concept and target indicating a "low" concept or viceversa). The main effect of Congruency was significant both considering participants (t1) or items (t2) as random factor ($t1 (19) = 3.63$, $p < .01$; $t2 (55) = 2.02$, $p = .04$) with faster RTs in the congruent than in the incongruent condition. In the analysis of the error rates, the main effect of Congruency was not significant ($ps > .22$).

META-ANALYSIS

In the meta-analysis we included all the studies that have explored the LCC from cue words presented in isolation (i.e., without context words). Table 1 shows a description of the methodological details of these studies. All the studies used a repeated measures design. Effect sizes (Cohen's d) were estimated from t statistics adopting the approach described by Morris and DeShon (2002) for repeated measures designs. The t value was estimated from the F value in two studies, Estes et al. (2008) and Gozli et al. (Exp3b, 2013). In order to estimate the t values of the critical comparisons in the three experiments reported by Verges and Duffy (2009), we contacted the authors who did not provide us with the required information. Therefore we could not include

these experiments in the meta-analysis. Sampling variance of the effect sizes were also estimated following Morris and DeShon (2002)[2].

Analyses were performed using the R package Metafor (Viechtbauer, 2010). The meta-analysis effect size was calculated by the restricted maximum-likelihood estimator. A random-effect model was adopted. The weighted mean of the effect size from the 15 studies was .17 with a 95% CI = [.002, .33], which can be considered a small effect (Cohen , 1988). Figure 1 shows a forest plot that summarizes the effect sizes of all the 15 studies. The analysis also showed that there is substantial and significant variability in the effect size across the studies, Q (14) = 34.47, p < .002, I^2 = 59.89%. In order to explore this heterogeneity, we compare the first model with a model containing the categorical predictor Response-Key alignment. This predictor has two values: horizontal vs. vertical. With horizontal alignment we mean that the response keys were placed on the same horizontal row in the keyboard (e.g., C & M). Whereas, with vertical alignment we intend response keys, which are horizontally dis-aligned, allowing for a vertical stretch (e.g., X & O). We used a maximum likelihood estimation (ML) method to perform this comparison. The analysis shows that the model including the moderator is also plausible (Bayes Factor = 1.3, with moderator's coefficient being QM (1) = 3.622, p= .057). In relation to the random-model, the weighted mean of the LCC effect size in the studies with vertical alignment increases substantially, .36 with a 95% CI = [.04 , .68], while in the studies with horizontal alignment decreases and was not significant, .07 with a 95% CI = [-.07, .22]. The studies with vertical alignment of the response keys are depicted in Red in Figure 1.

---

[2] Sampling variances were also estimated as described by Dunlap, Cortina, Vaslow & Burke (1996, Equation 4). Variances obtained with these two methods were almost identical (correlation of .99).

References

Bates, D., Maechler, M., & Bolker, B., 2011. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-39. <http://CRAN.R-project.org/package=lme4>.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1*, 170-177.

Goslin, J., Galluzzi, C., & Romani, C. (2014). PhonItalia: A phonological lexicon for Italian. Behavior Research Methods, 46, 872–886. doi: 10.3758/s13428-013-0400-8

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological methods*, *7*(1), 105.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raftery, A. E. (1995). Bayesian model selection in social research. Sociological Methodology, 25, 111-163.

Revelle, W. (2017) psych: Procedures for Personality and Psychological Research, Northwestern University,Evanston, Illinois, USA, https://CRAN.R-project.org/package=psych Version = 1.7.5

Schwarz, G. (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461–464. doi:10.1214/aos/1176344136

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. http://www.jstatsoft.org/v36/i03/.

DETAILED DESCRIPTION OF THE MATERIALS

Word stimuli used in Experiments 1, 2 and 3.

Top cue words: aereo (airplane), aquila (eagle), cielo (sky), gigante (giant), luna (moon), nuvola

(cloud), soffitto (ceiling), sole (sun), stella (star), tetto (roof), torre (tower), uccello (bird).

Bottom cue words: asfalto (asphalt), corallo (coral), erba (grass), formica (ant), pavimento (floor),

piede (foot), sabbia (sand), sasso (stone), scarpa (shoe), sommergibile (submarine), suolo

(ground), tappeto (rug).

Properties of the experimental words used in Experiments 1, 2 and 3.

	Top cue words			Bottom cue words		
	Mean	SD	Range	Mean	SD	Range
Top/Bottom Judgement	6.4	.4	5.9-6.9	2.2	.3	1.5-2.7
Frequency	133.8	146.2	15-514	62	49.1	2-159
Length	5.7	1.2	4-8	6.7	2.1	4-12
OrthUniq	5.3	1.1	4-7	6.2	2	4-11
Orth_N	6.3	4.9	1-17	3.4	2.8	0-9
Orth_N_MFreq	2.5	.6	1.5-3.5	2.6	1.3	0-3.8
Phon_N	4.9	3	1-10	3.6	2.8	1-9
Phon_N_MFreq	2.7	.9	1.5-5	2.8	1.1	.5-3.8

Note: Mean values, standard deviations (SD) and range values. Frequency = Colfis total frequency; Length = Number of letter; OrthUniq = Orthographic Uniqueness point; Orth_N = Orthographic neighborhood size; Orth_N_MFreq = Summed Orthographic neighborhood frequency; Phon_N = Phonological neighborhood size; Phon_N_MFreq = Summed Phonological neighborhood frequency. Frequency values were retrieved from the COLFIS database (Bertinetto, Burani, Laudanna, Marconi, Ratti, Rolando & Thornton, 1995). Orthographic and Phonological variables (i.e. OrthUniq, Orth_N, Orth_N_MFreq, Phon_N, Phon_N_MFreq) were retrieved from the PhonItalia database (Goslin, Galluzzi & Romani, 2014).

Word stimuli used in Experiment 4.

Top cue words: testa (head), cappello (hat), soffitto (ceiling), ramo (branch), aeroplano (airplane), lampadario (chandelier), condor (condor), bandiera (flag), rondine (swallow), cielo (sky), attico (loft), campanile (bell tower), aquilone (kite), satellite (satellite), antenna (antenna), luna (moon), chioma (foliage), nuvola (cloud), gabbiano (seagull), vetta (peak), torretta (turret), nido

(nest), razzo (rocket), sole (sun), tetto (roof), capelli (hair), camino (chimney), elicottero (helicopter), grattacielo (skyscraper), gru (crane), stelle (stars), funivia (cableway).

Bottom cue words: piedi (feet), scarpa (shoe), pavimento (floor), radice (root), sommergibile (submarine), tappeto (rug), verme (worm), binari (tracks), talpa (mole), terra (ground), cantina (cellar), marciapiede (sidewalk), canale (canal), tunnel (tunnel), strada (road), muschio (moss), seme (seed), pozzanghera (puddle), serpente (snake), fondale (backdrop), miniera (mine), tana (den), pozzo (well), tomba (tomb), fondamenta (foundation), pesce (fish), tombino (manhole), metropolitana (subway), cripta (crypt), rotaie (rails), lombrico (earthworm), relitto (wreckage).

Properties of the experimental words used in Experiment 4.

	Top cue words			Bottom cue words		
	Mean	SD	Range	Mean	SD	Range
Frequency	160.1	195.6	2-953	109.1	241.8	2-973
Length	6.7	2	3-11	6.9	2.3	4-13
OrthUniq	6.1	1.8	3-10	6.3	2.2	3-12
Orth_N	5.2	5.2	0-19	5.2	5.2	0-19
Orth_N_MFreq	2.2	1	0-3.9	2.2	1	0-3.9
Phon_N	4	4	0-15	4	4	0-15
Phon_N_MFreq	2.3	1.3	0-5	2.3	1.3	0-5

Note: For name values see previous note.

Word stimuli used in Experiments 6, 7, 8 and 10.

Top cue words: antenna (antenna), bandiera (flag), attico (loft), funivia (cableway), testa (head), lampadario (chandelier), nido (nest), torretta (turret), uccello (bird), aquilone (kite), campanile

(bell tower), gabbiano (seagull), condor (condor), elicottero (helicopter), razzo (rocket), rondine (swallow), aereo (airplane), aquila (eagle), grattacielo (skyscraper), nuvola (cloud), soffitto (ceiling), tetto (roof), cielo (sky), luna (moon), satellite (satellite), sole (sun), stella (star), vetta (peak).

Bottom cue words: fondale (backdrop), miniera (mine), pozzo (well), radice (root), sommergibile (submarine), cantina (cellar), cripta (crypt), erba (grass), fondamenta (foundation), pavimento (floor), pozzanghera (puddle), sabbia (sand), suolo (ground), verme (worm), asfalto (asphalt), rotaie (rails), serpente (snake), tappeto (rug), tombino (manhole), binari (tracks), corallo (coral), lombrico (earthworm), marciapiede (sidewalk), metropolitana (subway), piedi (feet), talpa (mole), sasso (stone), tana (den).

Filler words (not presented in Experiment 10): aceto (vinegar), agenzia (agency), anello (ring), bottiglia (bottle), bottone (button), bronzo (bronze), bullone (bolt), camice (coat), camicia (shirt), carta (paper), cartolina (postcard), cintura (belt), coltello (knife), crema (cream), cristallo (crystal), cucchiaio (spoon), cultura (culture), finestra (window), gesso (chalk), giornale (newspaper), inverno (winter), lama (blade), lega (league), legno (wood), lettera (letter), libro (book), macchia (spot), maglione (sweater), matita (pencil), messaggio (message), nome (name), numero (number), parola (word), pasta (pasta), pelle (leather), penna (pen), piatto (plate), quaderno (notebook), righello (ruler), risposta (answer), scatola (box),squadra (team), stanza (room), tabacco (tobacco), tegame (pan), tela (canvas), tempo (time), vassoio (tray), acquisto (purchase), gatto (cat), livello (level), magia (magic), nastro (ribbon), teatro (theatre), profumo (perfume), spilla (brooch

Properties of the experimental words used in Experiments 6, 7, 8 and 10.

	Top cue words			Bottom cue words		
	Mean	SD	Range	Mean	SD	Range

Top/Bottom Judgement	5.9	.4	5-6.9	1.8	.3	1.2-2.6
Frequency	104.1	200.8	2-953	53.4	107.3	2-578
Length	6.7	1.9	4-11	7.1	2.3	4-13
OrthUniq	6.1	1.6	3-10	6.5	2.2	4-12
Orth_N	5.1	5.4	0-19	3.6	3.5	0-14
Orth_N_MFreq	2.2	1	0-3.8	2.4	1	0-3.9
Phon_N	3.8	3.9	0-15	3.3	3.1	1-15
Phon_N_MFreq	2.2	1.2	0-5	2.4	1	.0-4.7

Note: For name values see previous note.

Word stimuli used in Experiment 9

Top cue-words: attic, buckle, cap, cloth, cloud, collar, curtains, ear, eye, handle, hat, laces, lamp, leaves, light, mane, peak, petal, roof, seat, snout, spout, stalk, steeple, surface, tower, tree, windows.

Bottom cue-words: base, boot, bristles, carpet, cellar, claw, coaster, corner, drain, drawer, floor, foot, hem, hooves, leg, lobby, milk, paw, pew, puddle, radio, roots, rug, seed, soles, stalk, stem, talon, tire, wheels.

Filler words: accusation, alcohol, algae, bacterium, bark, bee, bird, book, brake, chipmunk, clothes, competition, controversy, country, cow, cramp, current, equipment, extract, fertilizer, food, freshner, garden, itch, juice, lakes, language, law, lock, members, moisturizer, money, muffin, oil, powder, power, pressure, property, pulley, rash, report, riot, season, shape, shavings, shortage, show, slice, sound, soup, stain, stove, sugar, tension, town, treatment, turbine, virus, vote, wealth, wing, wreath.

Note: As in Estes et al. 2008, the words "cap" and "lamp" were repeated in the top cue condition;

the word "stalk" was repeated, once in the top cue condition and once in the bottom cue

condition; and the word "tower" was repeated, once in the top condition and once in the filler

condition.