From Distributional Structure to Meaning: Learning, Syntax, and the Emergence of Categorical and Graded Interpretations

Ling Sun (ls44@iu.edu)[1] **& Robert Goldstone**[2]
[1]Department of Linguistics, Cognitive Science Program, Indiana University
[2]Department of Psychological and Brain Sciences, Indiana University

Abstract

Adjectives often admit both categorical and graded interpretations, yet perceptual evidence alone typically underdetermines which interpretation is adopted during learning. We ask how learners converge on one interpretation and whether the learning task and syntactic structure play a causal role. Using novel adjectives mapped onto an arbitrary morph space with both discrete and continuous variance, we manipulate learning task (*Classification* vs. *Comparison*) and test generalization across syntactic frames in English and Mandarin. Learning task systematically shifts alignment between adjectives and perceptual dimensions. This alignment effect is amplified in Mandarin, where syntax overtly distinguishes categorical from graded predication. Critically, alignments do not reconfigure across post-learning tasks: category assignments remain stable, while graded information surfaces selectively in intensifier choice. Reaction-time asymmetries reveal directional markedness, with categorical-to-graded shifts incurring greater cost than the reverse. Together, the results show that adjectival interpretation emerges from the alignment between task structure and distributional variance, with syntax conditioning the strength of that alignment.

Keywords: categorization; concept learning; syntactic framing; gradability

Introduction

In linguistics, a long-standing distinction is drawn between gradable adjectives, which support comparison and degree modification, and non-gradable adjectives that resist such constructions (e.g., gradable: “very tall", “taller than" vs. non-gradable: “very wooden", “more wooden than"). This distinction is not merely grammatical but reflects differences in underlying conceptual structure (Gärdenfors, 2014; Paradis, 2001). Some adjectives encode meaning relative to an ordered scale of degrees (hereafter graded), whereas others denote discrete, category-defining properties (hereafter categorical) (Kennedy & McNally, 2005). Developmental evidence suggests that this duality poses genuine challenges for learners. Ryalls (2000) showed that children experience difficulty acquiring both the categorical and graded meaning of adjectives such as “tall". Similarly, Smith et al. (1992) found that when objects vary along multiple dimensions (e.g., shape and color), novel count nouns promote shape-based generalization, whereas adjectives yield more ambiguous patterns. Together, linguistic theory and acquisition studies suggest that adjectives exhibit a structurally dual organization, supporting both categorical and graded interpretations, and that learning or deploying these interpretations is neither trivial nor guaranteed.

More importantly, this body of work suggests that both learning task and syntactic framing play important roles in shaping how adjectival concepts are acquired. Ryalls (2000) showed that children acquire adjectives more readily through comparison alone than through mixed comparison and classification tasks, while Smith et al. (1992) demonstrated that syntactic category can bias learners’ attention toward particular dimensions (e.g., shape over color). What remains unaddressed, however, is how learners resolve the categorical-versus-graded duality of adjectival meaning when multiple interpretations are available, and whether learners flexibly recompose an adjective’s meaning on each use. In other words, current accounts do not explain how conceptual structure is selected and maintained when perceptual evidence alone underdetermines interpretation, which is the norm rather than the exception in language acquisition.

This question becomes tractable from a cross-linguistic perspective. Unlike English, where bare predication is compatible with both graded and categorical adjectives (“the chair is heavy", "the chair is wooden"), Mandarin Chinese encodes scale selection by syntax, such that bare predication typically licenses categorical adjectives (*zhe yizi shi mutou de* “the chair is wooden"), whereas graded adjectives require degree-marking constructions such as *hen* ‘very’ (Grano, 2012; Liu, 2010), making *zhe yizi shi zhong de* "the chair is heavy" odd. Crucially, this distinction goes beyond lexical category (noun vs. adjective) and directly constrains whether an adjective is interpreted categorically or in a graded manner (e.g., *shi xin de* "is new" implies a categorical interpretation, while *hen xin* "is new" yields a graded interpretation). A comparison between Mandarin and English therefore provides a natural testbed for examining how learning task and syntactic framing shape adjectival concept acquisition. If syntactic distinctions in a language shape what aspects of experience speakers habitually attend to when speaking (Slobin’s 1987 "Thinking for speaking" hypothesis), then we would expect that Mandarin speakers be more sensitive to the distributional distinction between categorical and graded dimensions.

In the present study, we draw on prior work demonstrating human sensitivity to structure in arbitrary, high-dimensional perceptual spaces (Goldstone & Steyvers, 2001; Hockema et al., 2005; Rogers et al., 2010). We use novel adjectives applied to arbitrary morph spaces that support two dimensional inter-

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

pretations. Participants learn novel adjectives under different learning tasks, classification ("it is A") versus comparison ("it is Aer than"), and their use of acquired adjectives was subsequently tested in bare predication ("it is A") and intensifier predication ("it is very A"). We ask whether the learning task led participants to rely on one dimension for adjective choice over the other, and whether Mandarin's obligatory syntactic marking of the categorical–graded distinction sharpens this bias relative to English.

We report three main findings. First, learning task influences which interpretation, categorical or graded, becomes stabilized for a novel adjective. Second, this effect is amplified in Mandarin, where syntactic structure constrains interpretation selection. Third, the resulting conceptual structure persists across contexts of use – changes in syntactic framing induce coercion via selective dimension separation, rather than reconfiguration or decay of the learned structure. Fourth, this coercion is asymmetric: shifting from a categorical to a graded construal incurs greater cost than the reverse, indicating directional markedness in adjectival representation. Together, these findings suggest that adjective meanings do not come with fixed conceptual structures. Instead, categorical and graded interpretations emerge through active alignment between learning task and the variational distribution of perceptual dimensions (discrete versus continuous), with syntax conditioning the alignment strength when perceptual information underdetermines interpretation.

Methodology

The online experiment employed a 2 (between-subject) × 2 (between-subject) × 2 (within-subject) mixed design. The two between-subjects factors were learning task (Classification vs. Comparison) and *Language* (English vs. Mandarin). The within-subjects factor was *Expression Type* (Bare predication vs. Intensifier predication).

Subjects A total of 64 Mandarin speakers and 87 English speakers (ages 18–35) were recruited. Participants in each language group were assigned to either *Classification* or *Comparison* Learning: Mandarin, $n = 32$ per condition; English, $n = 45$ and 43, respectively. All participants were L1 speakers and reported no visual or language impairments.

Exclusion criteria were designed to retain participants who either showed evidence of learning or maintained consistently strong performance throughout the task. Participants were retained if they demonstrated either (1) accuracy $\geq 70\%$ across the final 10 active-learning trials, accompanied by a positive learning slope, or (2) accuracy $\geq 70\%$ across both the first 10 and final 10 trials. These criteria excluded 18 participants, yielding a final sample of 132.

Stimuli Stimuli were drawn from a two-dimensional morph space defined by one binary variable (0 or 1, hereafter *Discrete*) and one continuous variable ranging from 0 to 1 (hereafter *Continuous*). Each stimulus was uniquely specified by a

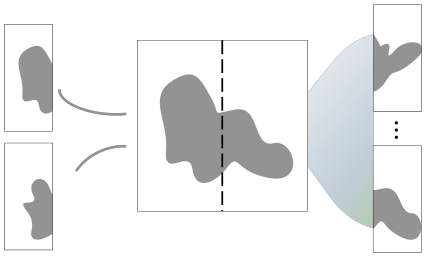

Figure 1: The arbitrary *Discrete* and *Continuous* dimension of Bézier-curve morphs.

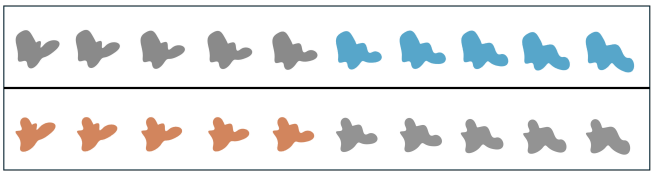

Figure 2: The *Discrete* and *Continuous* dimensions of the Bézier-curve morphs, with orange (A) and blue (B) indicating the two categories acquired during learning.

(*Discrete*, *Continuous*) coordinate – the left side of the stimulus morphed according to the *Discrete* variable, while the right side morphed according to the *Continuous* variable, as illustrated in Fig. 1.

Stimuli were labeled A if *Discrete* = 1 and *Continuous* < 0.5, and B if *Discrete* = 0 and *Continuous* > 0.5. This scheme ensured both dimensions were independently predictive during learning. As shown in Fig. 2, blue and yellow indicate the learned A and B regions, whereas grey indicates test-only incongruent cases where the *Discrete* and *Continuous* dimensions clash.

Adjectives A and B were novel words that participants were required to learn during the learning phase (*raflen* and *luprit* for English, *luzhi* and *machu* for Mandarin). The words were randomly assigned to adjectives A and B across participants to control for potential effects of prior linguistic associations.

Procedure The experiment consisted of three phases: baseline, learning, and testing.

- Baseline Phase (20 trials): The purpose of the baseline phase was to assess monotonic perception of the *Continuous* dimension. Participants were asked to position the 20 shapes, presented one at a time in random order, along a slider according to their similarity to the two endpoint shapes of the morph continuum. The endpoints corresponded to *Continuous* values of 0 and 1, randomly assigned to the left and right end of the slider, with the *Discrete* variable held constant at 1.
- Learning Phase (50 trials): Participants were trained on the novel adjective pairs (*raflen* and *luprit* for English; *luzhi* and *machu* for Mandarin). Learning consisted of interleaved passive and active trials and differed by learning task. In the

Classification Learning condition, passive trials presented each stimulus with a description (e.g., "The pink object is luprit"), and active trials required participants to select the correct adjective label. In the *Comparison Learning* condition, passive trials presented stimuli pairs with comparative descriptions (e.g., "The pink object is lupriter than the gray one"), and active trials required participants to select the correct comparative adjective.

Each session included 10 passive and 40 active trials. On active trials, participants could proceed only after selecting the correct response.

To discourage early convergence on a single perceptual dimension, a visual mask was applied sporadically across trials, randomly occluding one corner of the stimulus. Stimuli were evenly sampled from the labeled region of the morph space (i.e., the regions corresponding to adjectives A and B under the rule described in *Stimuli*).

- Testing Phase (48 trials): Participants completed two testing blocks of 24 trials each: a *bare predication* block and an *intensifier predication* block. Block order was counterbalanced to control for potential task-order effects.

  On each trial, participants were presented with a blank canvas and an incomplete sentence ("The pink object is __."). They were instructed to use the cursor to uncover the stimulus and complete the sentence using options defined by the block. In the *bare predication* block, they chose between the adjectival labels (e.g., "raflen" or "luprit"), whereas in the *intensifier predication* block, they selected from six options formed by crossing three intensifiers ("slightly," "somewhat," "very" in English; "youdian," "bijiao," "feichang" in Mandarin) with the two adjectives.

  Critically, both blocks contained incongruent trials, in which the *Continuous* and *Discrete* dimensions mapped onto different adjectives under the learned rule (e.g., *Continuous* > 0.5, *Discrete* = 1). Because the two dimensions provided conflicting cues, participants were forced to commit to one when applying the adjective.

  Participants were instructed to respond as quickly and accurately as possible. Both response choice and reaction time were recorded.

Results

Unless otherwise noted, all reported models are hierarchical Bayesian model fit to incongruent trials, where the *Discrete* and *Continuous* dimensions conflict and participants were forced to commit to one. We report three main models: (1) a logistic model predicting adjective choice (*Discrete*-aligned or not); (2) an ordinal model predicting intensifier strength ("slightly" $\prec$ "somewhat" $\prec$ "very"); and (3) a linear model predicting log response time. Effects from logistic and ordinal models are reported as odds ratios (OR); effects from linear models as posterior mean β. All credible intervals are 95% HDIs. An effect is credible when its HDI excludes the null: 1 for OR, 0 for β. Learning curves are analyzed using a generalized estimating equation (GEE; logit link, clustered on subject), which models population-average trends while accounting for within-subject correlation; effects are reported as OR with p-values.

Perceptual monotonicity and learning. Our results show that participants perceived the morph continuum monotonically, where stimuli with higher *Continuous* values were placed closer to the corresponding endpoint at the population level. Descriptively, accuracy was high across language and group conditions (English: $\mu_{Classification}$ = 86%; $\mu_{Comparison}$ = 82%; Mandarin: $\mu_{Classification}$ = 91%; $\mu_{Comparison}$ = 85%).

Participants also learned reliably over training, with comparable improvement across conditions. We fit a generalized estimating equation (GEE; logit link; clustering on subject) predicting trial-level correctness from trial order, learning group, and language. The probability of a correct response increased significantly over trials (OR = 7.31, $p < .001$), indicating robust learning. Although overall performance was higher in the *Classification* than *Comparison* group (OR = 1.59, $p = .017$) and higher for Mandarin than English speakers (OR = 1.56, $p = .021$), the interaction between learning group and trial order was not significant ($p = .88$). Thus, despite differences in overall task difficulty, learning rates were comparable across groups, with parallel improvements over training.

***Discrete* and *Continuous* evidence compete for adjective choice.** When the *Discrete* and *Continuous* dimensions conflicted, they directly competed. On incongruent trials, as the *Continuous* signal grew stronger, participants were less likely to choose the label associated with the *Discrete* dimension ($OR_{Continuous}$ = 0.54, HDI [0.43, 0.70]; Table 1, Model 1). For intensifier use, as the *Continuous* value became more extreme, creating greater conflict, participants were more likely to choose weaker intensifiers (e.g., "slightly") over stronger ones (e.g., "very") ($OR_{Continuous}$ = 0.69, HDI [0.57, 0.84]; Table 1, Model 2).

Response times further indicate that participants responded more rapidly when using stronger intensifiers. Model 3 (log(response time) ~ Intensifier + (1 + Intensifier | Subject)) shows that intensifier strength reliably predicted response latency. Selecting stronger intensifiers reliably predicted shorter response times (β = −0.048, HDI [−0.071, −0.026]), corresponding to an approximate 5% decrease in response time per unit increase in intensifier strength. This effect was relatively consistent across participants, with modest between-subject variability in slopes (slope SD = 0.08, HDI [0.05, 0.11]).

Learning shifts adjective cue dominance and reverses intensifier scaling. For adjective use, participants in the *Comparison* group were substantially less likely than those in the *Classification* group to choose the adjective associated with the *Discrete* dimension on incongruent trials ($OR_{Learning}$ = 0.28, HDI [0.14, 0.52]; Table 1, Model 1). This appears in Fig. 3 (top panel) as the red curve lying above the yellow one in both plots. However, the learning task did not reliably change

Table 1: Results of model 1 (adjective use; left columns) and model 2 (intensifier use; right column) on incongruent trials. $Y \sim$ Continuous + Learning + Continuous×Learning + (1 + Continuous | Subject). For model 1, Y is a binary outcome indicating response alignment with the *Discrete* dimension. For model 2, Y is an ordinal outcome indexing intensifier strength ("very" ≻ "somewhat" ≻ "slightly").

Subjects	*Adjective*			*Intensifier*		
	Predictor	OR	95% HDI	Predictor	OR	95% HDI
All	*Continuous*	0.54	[0.43, 0.70]	*Continuous*	0.69	[0.57, 0.84]
	Learning	0.28	[0.14, 0.52]	*Learning*	2.17	[1.34, 3.46]
	Continuous×Learning	0.75	[0.51, 1.08]	*Continuous×Learning*	3.71	[2.77, 4.93]
English	*Continuous*	0.46	[0.34, 0.63]	*Continuous*	0.82	[0.67, 1.00]
	Learning	0.52	[0.24, 1.11]	*Learning*	2.26	[1.46, 3.70]
	Continuous×Learning	0.92	[0.58, 1.48]	*Continuous×Learning*	2.79	[2.08, 3.85]
Mandarin	*Continuous*	0.61	[0.43, 0.84]	*Continuous*	0.55	[0.39, 0.81]
	Learning	0.17	[0.07, 0.46]	*Learning*	1.86	[0.83, 4.17]
	Continuous×Learning	0.70	[0.43, 1.14]	*Continuous×Learning*	4.96	[2.82, 8.52]

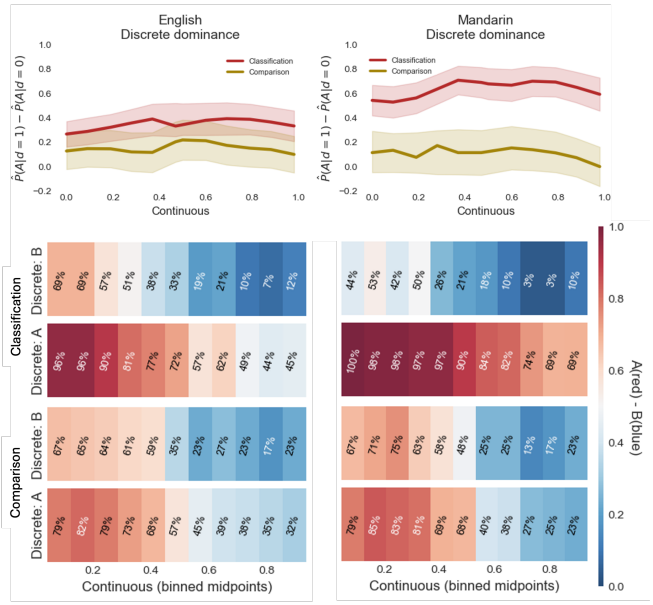

Figure 3: Top panels show the *Discrete* dominance index, $P(A|D{=}1) - P(A|D{=}0)$, as a function of the *Continuous* variable, with shaded 95% confidence intervals. Bottom panels show binned response proportions, with cell values indicating the percentage of trials on which adjective A was chosen (red = A; blue = B).

how this preference varied as a function of the *Continuous* signal ($OR_{Learning \times Continuous}$ = 0.75, HDI [0.51, 1.08]; Model 1), as both curves show a similar, slightly bowed, inverted-U shape. The same pattern is visible descriptively in Fig. 4 (bottom row), *Classification* learning encouraged alignment of the *Discrete* dimension with adjective use, as indicated by the response-percentage heatmap.

For intensifier use, the *Comparison* group overall used stronger intensifiers than the *Classification* group (Table 1, $OR_{Learning}$ = 2.17; Model 2). More importantly, the *Continuous* × *Learning* interaction revealed opposite effects of *Continuous* extremity across conditions ($OR_{Continuous \times Learning}$ = 3.71, HDI [2.77, 4.93]; Model 2). In the *Classification* group, more extreme *Continuous* values predicted weaker intensifiers ($OR_{Continuous}$ = 0.69, HDI [0.57, 0.84]; Model 2). This is illustrated by the red inverted-V-shaped dotted lines in both plots in the top panel of Fig. 4. In the *Comparison* group, more extreme *Continuous* values predicted stronger intensifiers ($OR_{Continuous}$ = 2.55, HDI [1.58, 4.12]; Model 2); this is illustrated by the yellow V-shaped dotted lines in both plots.

One interpretation is that the *Continuous* dimension is used to subdivide categories into "slightly," "somewhat," and "very" alternatives. The *Comparison* group aligned both intensifier and adjective with the *Continuous* dimension, yielding weaker intensifiers and a shift in adjective choice in the middle region. Thus, the learning task reshapes not only which dimension dominates but also how dimensions are functionally distributed across adjective and intensifier use.

Language amplifies the learning bias. For adjective use, the learning task effect on *Discrete* dominance was markedly stronger in Mandarin than in English. Among Mandarin speakers, *Classification* versus *Comparison* training produced a large shift in adjective-*Discrete* alignment ($OR_{Learning}$ = 0.17, HDI [0.07, 0.46]; Model 1); among English speakers, the corresponding shift was smaller and the HDI overlapped 1 ($OR_{Learning}$ = 0.52, HDI [0.24, 1.11]; Model 1). Crucially, the learning effect is not exclusive to Mandarin speakers. Adding Language and its interaction with the learning group to the main model yielded no credible language effects (Language: HDI [-1.16, 0.06]; Language × Group: HDI [-0.35, 1.31]). Most importantly, the main Learning effect remained robust ($OR_{Learning}$ = 0.22, HDI [0.10, 0.49]).

The learning task reliably shifted overall intensifier strength for English speakers ($OR_{Learning}$ = 2.26; Model 2), but not for Mandarin speakers ($OR_{Learning}$ = 1.86; Model 2). For English speakers at the continuous boundary (0.4 <*Continuous*<0.6),

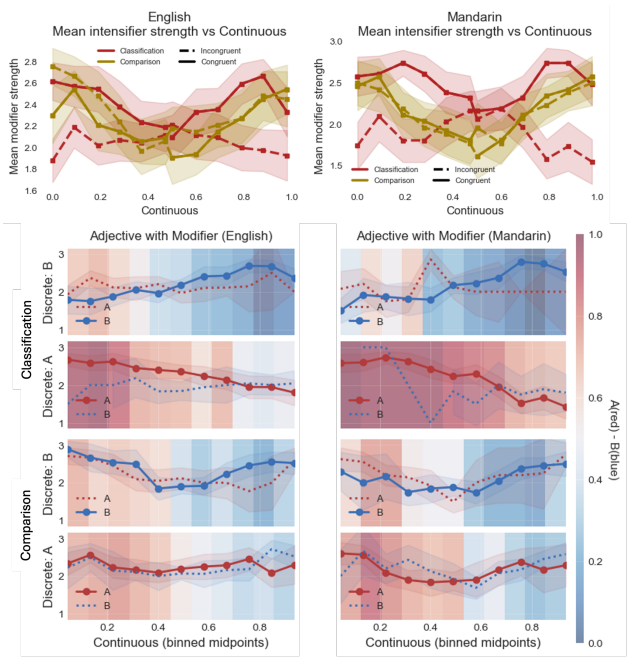

Figure 4: Top panels show mean intensifier strength as a function of the *Continuous* variable, with solid lines indicating congruent trials and dashed lines indicating incongruent trials; shaded ribbons represent 95% confidence intervals. Bottom panels show intensifier use overlaid with binned adjective proportions. Background cell shading reflects the percentage of trials in which adjective A was chosen (red = A; blue = B).

both learning groups preferred to respond "somewhat A/B," as shown by the converging yellow and red dotted curves (Fig. 4, top row). In contrast, the Mandarin *Comparison* group preferred "slightly A/B" (aligning both words with the *Continuous* dimension), whereas the *Classification* group preferred "somewhat A/B" (aligning the adjective with the *Discrete* dimension). Consequently, the Mandarin yellow and red dotted curves remain separated across this boundary, driving the stronger *Continuous* × *Learning* interaction. Together, these results indicate that language modulates how a learning task reorganizes dimensional alignment: English speakers showed a simple baseline shift, while Mandarin speakers exhibited a sharp reallocation of dimension roles across adjectives and intensifiers.

Asymmetric reaction-time costs across syntactic structures. Reaction-time patterns revealed an asymmetric processing cost: *Classification* training incurred difficulty in intensifier syntax, whereas *Comparison* training showed no comparable cost in bare predication. In the intensifier task, the *Comparison* group responded faster than the *Classification* group (β = -0.17, HDI [-0.36, -0.07], log(RT) ~ *Group* + (1 + *Intensifier* | *Subject*)), indicating increased processing cost when *Classification*-learned adjectives were used in degree-marked intensifier constructions. This RT shift was amplified for Mandarin speakers (β = 0.33, HDI [0.07, 0.59]) relative to English speakers (β = -0.11, HDI [-0.29, 0.06]). However, adding Language and its interaction with the learning group to the main model yielded no credible language effects (Language: HDI [-0.33, 0.08]; Language × Group: HDI [-0.08, 0.51]). No parallel slowing was observed when the *Comparison* group encountered bare predication across learning groups (HDI [-0.21, 0.06]).

Adjective choices, by contrast, remained stable regardless of whether participants completed the bare predication or intensifier block first. In both learning groups, syntactic task order did not reliably shift which dimension guided adjective choice. The hierarchical Bayesian logistic model (*Discrete-aligned* ~ *Continuous* + *Task Order* + *Continuous* × *Task Order* + (1 + *Continuous* | *Subject*)) showed no reliable order effects in either group: *Classification* — Task Order: HDI [0.54, 2.76], *Continuous* × Task Order: HDI [0.70, 1.48]; *Comparison* — Task Order: HDI [0.43, 2.66], *Continuous* × Task Order: HDI [0.62, 2.07]; all HDIs include 1.

Discussion

Learners map implicit distributional structure onto task form Prior work shows that learners use task and linguistic context to determine which perceptual dimensions matter. Adjectives like *tall* are learned more readily with only comparisons (Ryalls, 2000),count-nouns can bias attention toward shape (Smith et al., 1992). These findings demonstrate that word learning is shaped by the relation between language use and perceptual attention. However, they leave open a broader question: whether such effects are limited to particular dimensions, lexical categories, or familiar adjective meanings, or instead reflect a more general ability to align language task with perceptual structure.

Our findings reveal this more general principle. Learning task systematically determined which dimension guided adjective choice and how the remaining dimension was deployed in intensifier constructions. Under *Classification* learning, participants aligned adjective choice with the *Discrete* dimension; under *Comparison* learning, both adjective choice and intensifier strength aligned with the *Continuous* dimension. Learners did not simply encode associations; they actively aligned dimensional interpretation with task structures.

Crucially, this alignment was not driven by the inherent properties of particular dimensions, such as height being naturally graded, nor by lexical-category biases, such as count nouns directing attention to shape (Smith et al., 1992). The dimensions in our study were arbitrary; what mattered was their distributional structure: one learning task highlighted clustered structure consistent with category boundaries, whereas the other highlighted continuous variation consistent with gradation. More generally, our findings suggest that learners track how variation is distributed in the environment and align that distribution with the task structure. When the task emphasizes *Classification*, learners privilege dimensions that support discrete partitioning. When it emphasizes *Comparison*, learners

privilege dimensions that support graded evaluation.

This perspective helps explain why *tall* may be easier to acquire in comparative contexts (Ryalls, 2000): comparison highlights continuous variation, making it easier for learners to identify the relevant dimension. More generally, adjective learning appears to depend not only on which perceptual dimensions are available, but on how learners align the distributional structure of those dimensions with the conceptual/interpretative structure made relevant by language use.

Syntactic distinctions amplify learning bias The task–dimension alignment described above was further modulated by cross-linguistic differences in syntactic marking. This pattern is consistent with Slobin's "thinking for speaking" hypothesis, which holds that speakers attend to aspects of experience that their language routinely requires them to encode (Slobin, 1987).

Here, the relevant linguistic routine concerns adjectival predication: In English, bare predication licenses both categorical and graded readings. We routinely use the same frame ("This is X") whether X is graded like *tall* or categorical like *wooden*. Because English does not overtly distinguish these interpretations, learners do not attend to the difference between *Continuous* and *Discrete* dimensions as effectively. English speakers thus showed weaker alignment of this kind. Mandarin, by contrast, overtly encodes the distinction between categorical and graded interpretation. Bare predication strongly favors categorical readings, while graded interpretations require explicit degree marking. Accordingly, Mandarin speakers exhibited a substantially larger shift in *Discrete* dominance under *Classification* learning.

Syntax thus acts as an amplifier, as the "Thinking for speaking" hypothesis predicts: Mandarin speakers, who must habitually attend to the categorical–graded distinction, show stronger reorganization of dimensional roles than English speakers.Our findings extend "thinking for speaking" beyond previously studied domains by showing that syntactic routines can shape how learners align variance structure (*Continuous* vs. *Discrete*) with adjectival interpretation.

Our results suggest a straightforward account of Mandarin bare predication: it inherently favors categorical interpretations. This connects our findings with prior accounts of Mandarin adjectival predication (Grano, 2012; Liu, 2010), which note that gradable adjectives selectively appear in degree-marking constructions involving words like *hen* ("very"). Adjectives such as *new* or *full*, which readily allow for endpoint interpretations (*new* vs. *not new*; *full* vs. *not full*), easily fit the categorical bias imposed by bare predication. In contrast, adjectives like *tall*, which lack inherent partition points, are dispreferred in these bare contexts. This bias also explains why only intensifier-marked, but not bare, *new* can describe a second-hand book: bare predication forces a strict new/not-new partition that clashes with the item's second-hand status.

Stable concept structure with asymmetric coercion Our results favor a stable-alignment account over a fully flexible recomposition account. When multiple interpretations were available, learners did not appear to recompose adjective meaning anew for each syntactic frame. Instead, they maintained the adjective–dimension mapping established during learning, while syntax modulated how that mapping was coerced at test. If learners flexibly recomposed adjective meaning on each use, mismatched frames should have shifted adjective choices toward the *Continuous* dimension under degree marking and toward the *Discrete* dimension under bare predication. Instead, adjective choices remained stable: intensifier use could track the *Continuous* dimension while adjective choice remained *Discrete*-aligned, and *Continuous*-aligned choices persisted in bare predication.

Reaction times further suggest asymmetric coercion. Categorical-to-graded shifts incurred reliable slowing, whereas graded-to-categorical shifts did not. In other words, graded-to-categorical shift was relatively easy, whereas categorical-to-graded shift was costly. This directional asymmetry aligns with broader developmental patterns. Entity-based concepts and simple predication are acquired earlier than relational meanings and degree constructions (e.g., Gentner, 2005). Following Gentner (2005), categorical interpretations resemble one-place predicates, whereas graded interpretations encode relational structure over an ordered dimension, making graded construal demand structure beyond what categorical representations supply.

Conclusion and limitations

This study shows that categorical and graded interpretations are not fixed properties of adjectives, nor are they determined solely by perceptual input. Instead, they emerge from the alignment between distributional structure in the environment and the structure of the learning task. Syntax modulates the strength of this alignment, amplifying learning biases in languages that overtly distinguish categorical and graded predication. Crucially, once established, these alignments remain stable across contexts of use: syntactic coercion selectively recruits secondary dimensions rather than reorganizing category structure. Together, these findings suggest that adjectival meaning reflects learning-stage alignment between variance structure and linguistic form, not inherent scale properties or online reanalysis.

Our conclusions are based on novel adjectives mapped onto arbitrary dimensions, allowing us to isolate distributional structure from lexical semantics. Whether the same alignment mechanisms operate for entrenched natural adjectives remains to be tested. In addition, although we demonstrate cross-linguistic amplification in Mandarin and English, broader typological comparison is needed to determine how general syntactic amplification is.

Acknowledgments

The research was conducted in accordance with the ethical standards of the Indiana University Human Research Protection Program (HRPP). The study protocol was reviewed and approved by the Indiana University Institutional Review Board (Protocol #: 25337).

AI tools were used to assist with programming tasks related to experimental implementation and statistical coding. All experimental design, hypotheses, analysis plans, codebooks, data validation procedures, and interpretation of results were developed and verified by the authors. AI tools were not used for study design, theoretical development, data analysis decisions, or result interpretation.

References

Gärdenfors, P. (2014, January). *The geometry of meaning: Semantics based on conceptual spaces*. The MIT Press. https://doi.org/10.7551/mitpress/9629.001.0001

Gentner, D. (2005). The development of relational category knowledge. In D. H. Rakison & L. M. Oakes (Eds.), *Building object categories in developmental time* (pp. 245–275). Lawrence Erlbaum Associates.

Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of experimental psychology: General*, *130*(1), 116.

Grano, T. (2012). Mandarin *hen* and Universal Markedness in Gradable Adjectives. *Natural Language & Linguistic Theory*, *30*(2), 513–565.

Hockema, S. A., Blair, M. R., & Goldstone, R. L. (2005). Differentiation for novel dimensions. *Proceedings of the twenty-seventh annual conference of the cognitive science society*, 953–958.

Kennedy, C., & McNally, L. (2005). Scale Structure, Degree Modification, and the Semantics of Gradable Predicates. *Language*, *81*, 345–381.

Liu, C.-S. L. (2010). The Positive Morpheme in Chinese and the Adjectival Structure. *Lingua*, *120*(4), 1010–1056.

Paradis, C. (2001). Adjectives and boundedness. *Cognitive linguistics*, *12*, 47–64.

Rogers, T., Kalish, C., Harrison, J., Zhu, J., & Gibson, B. (2010). Humans learn using manifolds, reluctantly. *Advances in neural information processing systems*, *23*.

Ryalls, B. O. (2000). Dimensional adjectives: Factors affecting children's ability to compare objects using novel words. *Journal of Experimental Child Psychology*, *76*(1), 26–49.

Slobin, D. I. (1987). Thinking for speaking. *Annual Meeting of the Berkeley Linguistics Society*, 435–445.

Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Developmental Psychology*, *28*(2), 273.