

Language induced changes of mind: shifts in categorical colour perception after a short language learning experience

Morgane Peirollo^{a,b,*,1}, Amie Fairs^a, Guillaume Thierry^c and Elin Runnqvist^{a,b}

^aAix-Marseille Université, CNRS, LPL, 5, avenue Pasteur, Aix-en-Provence, 13100, France

^bInstitute of Language, Communication and the Brain, 5, avenue Pasteur, Aix-en-Provence, 13100, France

^cBangor University, Gwynedd, LL57 2DG, Bangor, United Kingdom

ARTICLE INFO

Keywords:

language learning
bilingual language production
lexico-semantic competition
grounded learning
colour perception
categorical perception
reorganization of semantic space

ABSTRACT

This study investigated the reorganization of perceptual categories following a short language learning experience. A first goal was to investigate the interaction of previous labels with learning and potential language-induced effects on categorical perception. A second goal was to determine whether language learning might make similar colours more distinctly perceived (within category separation) and/or more dissimilar colours less distinctly perceived (between category compression). A third goal was to assess whether potential effects would be persistent, indicating authentic learning and reorganization of the semantic space. In two experiments, native French participants were asked to learn new labels for two nuances of each of four hues. Across participants, we manipulated whether the colours had a previous robust label (ETN colours) or not (HTN colours). Linguistic similarity of labels for different nuances and hues was manipulated within participants. Speakers then had to rate the perceived similarity of colour pairs. Results showed that colours that had a previous robust label were persistently rated as more similar if they had similar new labels. These findings highlight 1) a key role of spreading activation and lexico-semantic strengthening induced by competition as a drive of second language learning, and 2) an impressive malleability of the human perception-based interpretation of the world.

1. Introduction

1.1. Background literature

A large body of work has shown that categorical perception is sustained and affected by language: for example, two objects sharing a label are perceived as belonging together, compared to objects without labels (Lupyan, Abdel Rahman, Boroditsky and Clark (2020); Gervits, Johanson and Papafragou (2016)). In one study, Johanson and Papafragou (2016) showed that perceptually similar objects are more likely to be grouped together if they have similar labels. This evidence suggests that linguistic labels applied to objects can affect the non-linguistic perception of these objects.

Following those findings, studies across languages that have a more or less detailed vocabulary related to a given perceptual domain have illustrated the effects of long-term linguistic influences on our way of categorizing (Thierry, Athanasopoulos, Wiggett, Dering and Kuipers (2009); Athanasopoulos, Dering, Wiggett, Kuipers and Thierry (2010); Winawer, Witthoft, Frank, Wu, Wade and Boroditsky (2007)). For example, Greek has two different words to refer to dark and light blue (respectively *ble* and *ghalazio*) while English only has one (*blue*). It has been shown that Greek speakers rate light and dark blue as being less similar compared to English speakers (e.g., Thierry et al. (2009)). Furthermore, besides differences in perceived similarity, Greek speakers also showed a larger modulation of the visual MisMatch negativity (an Event Related Potential, ERP, component) compared to English speakers in an oddball task that used one of the blue nuances as standard (e.g., light blue) and the other one as deviant (e.g., dark blue). The oddball task is a procedure designed to investigate early perceptual processes indexed through ERPs. Sequences of identical stimuli are presented and randomly interrupted by a deviant stimulus, or oddball. Previous studies have shown that the visual MisMatch Negativity (vMMN) is known to be a reliable marker of pre-attentive cerebral detection of visual stimuli (e.g., Stefanics, Astikainen and Czigler (2015)). Hence, it would seem that the perceptual differences related

*Corresponding author

✉ morgane.peirollo@mpi.nl (M. Peirollo)

to having one or two labels for different nuances of a colour have a pre-attentive origin. Interestingly, another study showed that these language-related differences in perception can change over time as a consequence of linguistic immersion. Concretely, Athanasopoulos et al. (2010) showed that Greek speakers that had spent a long time in the UK (average of 42.6 months) showed a reduced vMMN for different nuances of blue compared to Greek speakers who had only spent a short period in the UK (average of 7.2 months). Along the same lines, Juárez, Sicotte, Thériault and Harnad (2019) showed that when participants learn to associate two distinct labels to sort multi-featured visual textures into two categories, perceptual similarities are reinforced (within category compression) and perceptual differences are enhanced (between category separation) compared to when no label was given. This effect was revealed both in dissimilarity ratings and on the early perception-related (N1) and late decision-related (LPC) ERP components. Taken together, these findings suggest that the relationship between categorical perception and language is rather malleable. However, several important open questions remain concerning the nature and boundaries of such language-induced changes of mind.

1.2. The present study

A first question is whether any effect on categorical perception would be observed when learning a new label for an object that already has a label in the native language. The study of Thierry et al. (2009) suggests that long-term linguistic-induced effects on perception can be observed. However, given the timescale of these effects, it is difficult to know whether they relate to learning a new language or forgetting the native language (i.e., first language attrition, e.g., Schmid (2013); Runnqvist and Costa (2012)). In turn, Juárez Juárez et al. (2019) focused on learning labels for objects that did not have any previous label. Therefore, it is yet unknown how the dynamics of learning and the categorization of reality would be affected by a short learning session teaching distinctions that differ from those of the native language of the participants, for objects that already have labels in the participant's mother tongue.

To understand why having or not having a previous label might be relevant for the effects of interest here, we need to briefly turn to some key principles of bilingual language production, namely co-activation of languages and language control. A large body of work has focused on speech dynamics in the context of bilingualism, showing that both languages are activated even when a task is performed in only one of them (e.g., Colomé (2001); Wu and Thierry (2010); Kroll, Bobb, Misra and Guo (2008)). The Revised Hierarchical Model (RHM) developed by Kroll and Stewart (1994) has been proposed to explain asymmetries found between first language and late-acquired second language in the access of the semantic network. According to this model, the first language (L1) has direct access to semantic components, whereas the second language (L2) requires translation into the first language to access the meaning of a word. According to this theory, co-activation of L1 when speaking in L2 would thus be the result of an automatic translation in the case of low proficient speakers or second language learners. For more proficient speakers, the co-activation would rather be due to automatic spreading activation (e.g., Wu and Thierry (2010); Runnqvist, Strijkers, Alario and Costa (2012)). Co-activation is thought to be dependent on the strength of lexical representations, leading words in L1 to be more co-active during L2 production than vice versa during the initial phases of learning, or for low proficient bilingual speakers (e.g., Runnqvist, Strijkers and Costa (2019)). Given that a unique lexical candidate has to be selected for production, such co-activation from the language currently not in use is a potential source of interference. There is no consensus in the literature concerning how exactly speakers cope with the potentially interfering co-activation from the non-target language when speaking. However, two influential alternatives of bilingual language control have been put forward. The first alternative is that co-activated lexical items from the non-target language have to be inhibited in order for a speaker to select the appropriate word in the desired language (e.g., Green (1998)). The second alternative is that lexical items from the weaker language undergo disproportional strengthening in their lexico-semantic mappings in order to facilitate their production (e.g., Runnqvist et al. (2019)). While this kind of bilingual language control is thought to be in place when speaking in a second language, little is known about what role it plays, if any, during the process of learning. These complex interactions between the co-activation of previously existing labels and bilingual language control are likely to have an impact on language learning induced effects of categorical perception. Returning to the present study, most if not all languages count on a reduced set of words to name the prototypical colours such as blue, purple, red, orange, yellow, and green. The lexical frequency of these labels tend to be extremely high, and hence they are conceivably highly co-activated when learning a new label for these colours. Depending on how this co-activation is dealt with, it may be expected to impact categorical perception differently: if speakers rely on the native language labels to access semantics through automatic translation as proposed in the RHM, interference from these can be expected to be maximal, potentially cancelling out any effect of the newly learnt labels

on categorical perception. However, this strength in coactivation can also be expected to interact with processes of bilingual language control. In order to learn a new label for a colour that has a previous label very strongly associated to it, presumably speakers need to either inhibit the native language label very strongly, or strengthen the new label very strongly (e.g., Runnqvist et al. (2019)). This interaction with bilingual language control, especially if it is carried out by strengthening the weak labels that are being learnt, could lead the labels of colours with a previously strong label attached to them (blue, red, etc.) to have a large impact on categorical perception.

A second question is whether the effects of learning a new language on categorical perception would extend to within category separation (perceptual differences are reinforced) and between category compression (perceptual similarities are enhanced). The within category separation effects observed so far in the literature only concerned comparisons between a distinction present in the native language of some participants and absent in the native language of others (i.e., *ble* and *ghalazio* in Greek versus *blue* in English). The effects of a language learning experience on this distinction was in the sense of within category compression (i.e., a loss of the distinction). Concerning between category compression (i.e., increased similarity of objects belonging to different categories such as blue and green), to the best of our knowledge no previous study has addressed this phenomenon. Nevertheless, it is perhaps the most interesting condition to attain a better understanding of the pure impact of language on categorical perception, as it is the only one that completely dissociates perceptual similarity and linguistic similarity.

A third question investigates whether any potential language-induced effects on categorical perception obtained in a single experimental session would reflect actual persistent lexical learning and reorganization of perceptual categories. Alternatively, it might be the result of a transient stimulus response association without any consequences for language or categorical perception beyond the experimental session. Novel stimulus response associations created within a single experimental session have indeed been shown to impact perception by, for instance, reducing the conflict adaptation in Stroop like tasks (e.g., Braem, Verguts and Notebaert (2011); Raz, Kirsch, Pollard and Nitkin-Kaner (2006)). In the context of word learning, the stimulus response association could be assimilated to the rapid, initial familiarization phase proposed by Davis and Gaskell (2009) in the adapted model of Complementary Learning Systems (CLS), which would be followed by a slow lexical consolidation. First, the word forms an episodic memory trace in the hippocampal system. Then, the word goes through a period of offline consolidation before being integrated in neocortical structures. To sum up, the use of multiple sessions seems appropriate to assess whether new labels have become lexicalized, arguably leading to a reorganization of the semantic space (in the present study, of perceptual categories).

In the current study, we aimed at answering the three questions outlined above through two experiments. In both experiments, participants learnt to associate labels to a series of colours. Crucially, certain colours were given the same or closely resembling labels (e.g., *djols* for blue and purple) while other colours were given labels without any phonological resemblance (e.g., *tsalp* for yellow and *snild* for orange). Afterwards, participants' categorical perception of the colours was assessed by having them rate the similarity between pairs of colours. An effect of language on categorical perception would be reflected through modulations of the similarity ratings by the similarity of the linguistic labels of a given pair of colours. To address the first question outlined above (i.e., will effects on categorical perception be observed when learning a new label for an object that already has a label in the native language?), we introduced a manipulation related to the robustness of previous linguistic labels associated to the colours. For half of the participants in both experiments, we used *Easy-to-name* colours (ETN), i.e., colours with a strongly attached label (e.g., red, green, blue) for which the speakers are close to full agreement when naming. For the other half we used *Hard-to-name* colours (HTN), i.e., colours with low naming agreement (e.g., complex shades such as taupe, turquoise, salmon) that possess no stable label since speakers tend to feel uncertain about their colour names. We reasoned that the previous native (henceforth L1) ETN colour names would be more strongly coactivated during learning of the new labels (henceforth LNew) compared to L1 labels for HTN colours. If maintaining this co-activation is necessary for the learning of the new label to take place as proposed by the RHM, one might expect that the more co-active the previous labels, the less impact of the new labels on categorical perception. That is, learning new labels for ETN colours should result in a much smaller effect on categorical perception compared to learning new labels for HTN colours. If the co-activation is countered by inhibiting the previously existing labels, the effects on categorical perception can be expected to be similar between the ETN and the HTN conditions. Finally, if the co-activation is countered by a proportional strengthening of the new labels for the colours, learning new labels for ETN colours should result in a much larger effect on categorical perception compared to learning new labels for HTN colours. To address the second question outlined above (i.e., will the effects of learning a new language on categorical perception extend to within category separation and between category compression?) we manipulated two variables: Label (same/ shared root/ different in Experiment 1 and same/ different in Experiment 2) and Category (same -e.g., light blue and dark blue- or different -e.g., blue and purple).

Contrasting ratings for colour pairs of the same category that had been given the same or different labels would allow us to assess within category separation (e.g., shades of blue perceived as more distinct after learning to associate them to dissimilar labels). Contrasting ratings for colour pairs of different categories that had been given the same or different labels would allow us to assess between category compression (e.g., blue and purple perceived as less distinct after learning to associate them to similar labels). Finally, to address the third question outlined above (i.e., would language-induced effects on categorical perception obtained in the context of a single experimental session reflect actual persistent lexical learning and reorganization of perceptual categories or only a transient stimulus response association?), in Experiment 1 we introduced a second experimental session taking place one day after the first one. In this session, participants completed the same similarity ratings as on the previous day. Observing effects of categorical perception on the second session would suggest that authentic learning and reorganization of perceptual categories had taken place. This is because on the second day, the ratings and oddball tasks were not immediately preceded by the learning, precluding transient stimulus response associations to account for any effects.

2. Experiment 1

Experiment 1 investigated the effects of learning a new vocabulary on categorical perception of colours. Participants carried out a learning phase where they had to associate newly-learned pseudowords to 8 colours (four different hues - blue, purple, orange, yellow - with two nuances each - light and dark). Half of the nuances of the same hue (e.g., light and dark blue, light and dark yellow) were associated with one label and the other half with two labels (e.g., light and dark purple, light and dark orange). When two labels were associated to the nuances of the same hue, the labels shared the root but differed in the coda (e.g., *djols* and *djoft* for dark and light blue). While nuances of different categories never shared a label, half of them also shared the root (e.g., *djonz* for purple). Table 1 illustrates all the stimuli and experimental conditions of Experiment 1. After the learning task, participants rated the similarity of all possible combinations of the colours. If the newly learnt colour vocabulary would have an effect on categorical perception, nuances of the same hue receiving two labels would be rated as perceptually less similar than colours receiving only one label (i.e., within category separation), and nuances belonging to different hues that shared a root would be rated as more similar than those that had entirely different labels (between category compression). Furthermore, to assess the potential interaction between the robustness of previously existing labels (leading to a varying amount of co-activation) and the effects on categorical perception induced by the newly learnt labels, participants were randomly assigned to one of two different groups, tested on distinct colour sets: 1) Easy To Name colours (ETN), expected to result in strong co-activation; 2) Hard To Name colours (HTN), expected to result in a weak co-activation. Depending on whether the co-activation of previous labels is maintained, inhibited or leads to a proportional strengthening of the new labels, one would expect larger effects on categorical perception for the HTN, no difference between ETN and HTN or larger effects for ETN respectively. Finally, the duration of the effects on categorical perception were assessed through a second session carried out 24 hours later in which participants performed the same similarity ratings as on the previous day, and then they were asked to name the colours to assess the retention rate of the newly learnt labels. If the potential effects on perpetual categorization reflect persistent learning, they would be expected to be present also on the second session.

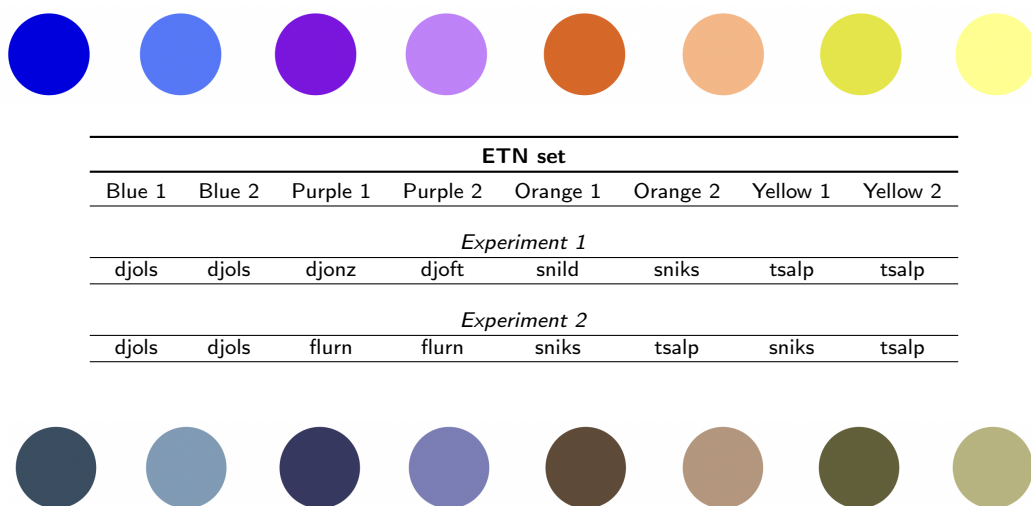
2.1. Material and methods

Participants. 72 participants gave written consent to take part in the experiment ($M=24.19$ years; $SD=3.19$; [19;32]; 60 female, 12 male). All participants were native French-speakers recruited through Aix-Marseille University and compensated 15 euros via PayPal. The experiment took an average of 55 minutes to be completed, spread across two sessions (45 minutes for the first session, 10 minutes for the second session). All participants reported no hearing, speech or neurological disorders, and had normal or corrected-to-normal vision, with no colour perception disorders. Data from two participants were rejected due to equipment failure (microphone issues, one in each group), and data from four participants were rejected due to a misunderstanding of the task (one in the ETN group, three in the HTN group), leaving a sample of 66 participants data for analysis (34 in ETN, 32 in HTN). In order to control for language facilitation effects (Rolstad and MacSwan (2014); Incera and McLennan (2018)), participants were asked to self assess their language proficiency through a preliminary questionnaire. The scores representing the proficiency, age of onset and frequency of use are provided in the Supplementary Material.

Stimuli

Colours. Stimuli were composed of 16 computer-simulated colour chips, 8 belonging to the ETN category and 8 to the

Language induced changes of mind



ETN set							
Blue 1	Blue 2	Purple 1	Purple 2	Orange 1	Orange 2	Yellow 1	Yellow 2
Experiment 1							
djols	djols	djonz	djoft	snild	sniks	tsalp	tsalp
Experiment 2							
djols	djols	flurn	flurn	sniks	tsalp	sniks	tsalp

HTN set							
Grey 1	Grey 2	Lilac 1	Lilac 2	Maroon 1	Maroon 2	Gold 1	Gold 2

Table 1
ETN & HTN colour sets and labels used in Experiments 1 and 2.

HTN category. The variable Colour category was divided between Within and Across category. The Within category was composed of a pair of a dark and light nuance of the same colour, created by contrasting luminance. The Across category was composed of colours contrasted according to their hue. The colour sets were pre-tested to ensure that the perceptual differences between ETN and HTN colour sets were significant. Colour sets are displayed in Table 1 and detailed information about the full colour selection procedure is given in the Supplementary Material.

Pseudowords. Seven speech units conforming to the phonotactic rules of French while being devoid of meaning (i.e., pseudowords) were created (Dell (1995)). According to the work of Johnson and Eisler (2012) and Lima and Inhoff (1985) on the importance of the first letter in word recognition, the pseudowords were created by eliminating each of the initial phonemes used in common colour names in seven languages (French, English, Spanish, Italian, German, Arabic and Portuguese), chosen because these are common languages spoken in France ¹. To match the learning difficulty of the pseudowords, length and syllabic structure were held constant. To control for potential facilitation effects, each pseudoword was checked for lexical neighbours (Duyck, Desmet, Verbeke and Brysbaert (2004)). Finally, the set of pseudowords was pre-tested to ensure equal perceived pronunciation difficulty and thus determine suitability of the stimuli set. Detailed information about the complete pseudoword selection procedure is accessible in Additional Material.

2.2. Instrumentation and measures

Experimental design. Four variables of experimental interest were manipulated. The first variable was Label, which refers to the label given of a given colour pair (see table 1 for examples of colour sets). Nuances belonging to the same hue (e.g., light and dark blue) could either receive the same label (*djols*) or two labels with the same root but a different coda (*djols*, *djonz*). Nuances belonging to different hues (e.g., blue and purple) could either receive labels that shared the root and differed in the coda (*djols*, *djonz*), or labels that were entirely different (*djols*, *sniks*). The second variable was Colour category, which refers to the type of colours compared in a colour pair (two levels: across (e.g., blue and purple); within (e.g., dark blue and light blue)); The third variable was Group, which refers to the easy or hard to name colour sets (two levels: ETN, HTN). Finally, the fourth variable was Session number, referring to the first or second day of experiment (two levels: session one, session two).

¹<https://ec.europa.eu/eurostat/web/microdata/adult-education-survey>

The dependent variables were naming accuracy for the learning task and the similarity ratings for the rating task. To avoid any stimulus effects, all colour-label associations were fully randomized across participants in both groups. Due to a bug in the experiment, the experimental conditions were not perfectly randomized, resulting in one condition being tested five times instead of one. Specifically, the ETN group only displayed 49 colour pairs in the similarity rating task, instead of 56. This resulted in the Within category colour pair Purple 1-Purple 2 to be presented only once; the three remaining colour pairs were Blue 1-Yellow 1, Blue 2-Yellow 2 and Orange 2-Purple 2, which are colour pairs belonging to the baseline category (i.e., colours that did not belong to the categories of interest, for example Blue 1 and Orange 2). Post-hoc analyses showed that the results were not impacted by the imbalance ($\chi = 3.089$; $p = 0.797$). The experiment was created and run online on FindingFive².

Procedure.

Naming out loud task. The naming out loud task was divided into learning blocks and testing blocks. In the learning blocks, the eight label-colour associations (e.g., Blue 1 was called *djols*) were presented five times each in a randomized order in the ten experimental blocks, so a total of 400 learning trials. For each colour, the label colour pair was presented for 1000 ms, followed by the colour chip for 3000 ms, during which the participant was instructed to name out loud the label associated with the colour displayed (e.g., if they saw Blue 1 they had to name out loud *djols*). A blank screen of 700 ms intervened between each trial. At the end of each learning block, a testing block was presented with all 8 colour chips displayed one time each for 3000 ms, during which the participant was instructed to name out loud the label associated with the colour displayed as quickly and accurately as possible. The testing trials were composed of a total of 80 stimuli, with all stimuli presented once in a row at the end of every experimental block. Each participant completed ten blocks of the task. The order of stimuli within the blocks was randomized across and within participants.

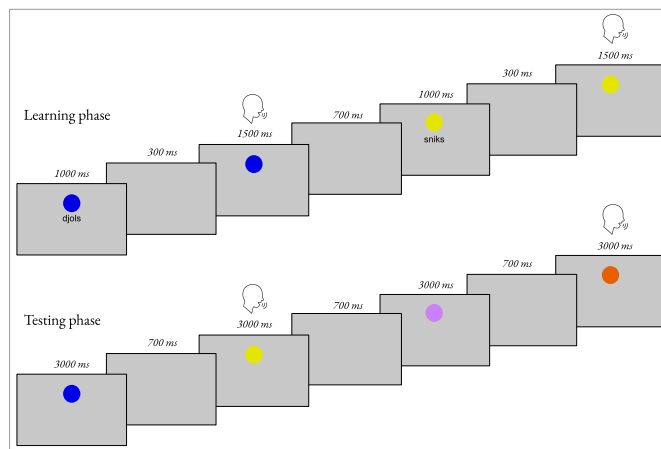


Figure 1: Naming out loud sequential procedure

Similarity rating task. After completing the naming out loud trials, participants were tested in a similarity rating task. Participants were shown all colour pair combinations that were possible by randomly combining the colours previously seen during the first stage of the experiment. The 28 possible colour pairs (e.g., Blue 1 and Purple 2) were all presented twice each by counterbalancing the order of appearance (i.e., Blue 1 Purple 1; Purple 1 Blue 1), to a total of 56 trials. Participants were asked to rate the similarity of the colour pairs on a scale of 1 to 6 (1: very different, 6: very similar) with a button-press response on the numeric keyboard. They were instructed to make all judgments as quickly and accurately as possible. Once a rating was entered, the screen was cleared for 500 ms before the next colour pair appeared.

Data preprocessing. Each trial belonging to the testing blocks of the naming out loud task was manually checked to determine the error rate. Naming accuracy in the testing blocks was analysed to determine the extent to which participants had learnt the colour-label associations. Naming accuracy of the first session was calculated over experimental blocks one to ten. Naming accuracy of the second session (retention rate) was calculated over the eleventh

²European server on <https://eu.findingfive.com/>

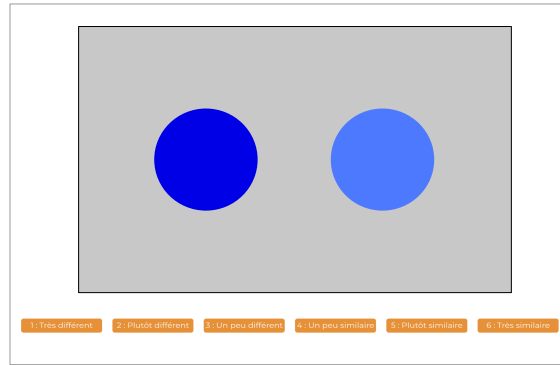


Figure 2: Similarity rating procedure.

experimental block. In order to remove outliers, trials in the similarity ratings with response latencies more than 2 standard deviations above or below the participant's mean were not considered in the analyses.

Statistical analysis. The data were analysed using the lme4 package (Bates, Mächler, Bolker and Walker (2015)) in R version 3.2.2 (R Core Team (2020)). Naming accuracy was analysed using generalized linear mixed models (GLMM) with a binomial link function (e.g., Jaeger (2008)), estimating the conditional probability of a response given the random effects and covariate values. Similarity ratings were analysed with linear mixed models (LMM), estimating the influence of fixed and random covariates on the response (e.g., Baayen, Davidson and Bates (2008)). In all models, Participants and Items were included as crossed random effects (i.e., intercept estimates), allowing to tease apart the influence of subjects/items on their repeated observations from the influence of the fixed effects of theoretical interest. Naming accuracy and similarity ratings were fitted in separate models. Naming accuracy was modelled to assess the learning effect over the testing blocks in the first session and to assess the retention rate after the second session. Session one was compared to session two by comparing the accuracy of the tenth block (last block of session one) and the accuracy of the only block of the second session. These models included the fixed covariate Trial number, and the main factors of Block number (1 to 11; reverse helmert coded) and Group (sum to zero coded). Similarity ratings were modelled to assess the impact of learning a new linguistic label on categorical perception of colours belonging to the same hue or different hues. The similarity rating was the dependent variable and was centred. The fixed effects were Label category, Group and Session number. Trial number was defined as a covariate. The continuous predictor Trial number was centred and scaled to reduce skew and help with model fit. Group and Session number were sum to zero contrast coded. Label was specially coded to address the comparisons of interest, i.e., A = same label same hue, B = shared root same hue, C = shared root different hue, D = different label different hue, C* = shared root different hue different luminance, D* = different label different hue different luminance, E = baseline. The crucial contrasts to assess our hypotheses concerning within category separation and between category compression were A versus B and C versus D, respectively. Additional contrasts were carried out to assess the basic perceptual effects of differences in luminance (C versus C*; D versus D*) and hue sensitivity (AB versus E). A placeholder contrast was also created (ABE versus CC*DD*) to ensure model robustness (the models are available in the Additional Material).

Furthermore, the interactions between the contrasts of interest and Group (ETN and HTN) as well as Session (one and two) were assessed to address our hypotheses related to the lexical robustness of previous labels and persistent learning versus transient stimulus response mappings. For all models, a t-value greater than $|1.92|$ was taken as a marker of significance.

2.3. Results

Naming accuracy. Mean naming accuracy is reported in Figure 3, and the statistical models are summarized in Table 3. We observed no main effect of Group ($t = -0.589$) showing that ETN participants did not perform better than HTN in the naming out loud task, but a main effect of session was present ($t = 2.004$), with an average lower error rate in Session 2 compared to Session 1. No interaction between Group and Block number was found for any of the experimental blocks, but a main effect of Block number was found for each block. This shows that 1) the performances

Language induced changes of mind

Colour pairs	Labels	Label category	Comparison levels	Category
Blue 1 Blue 2	djols djols	same label	A versus B	Within category
Blue 1 Blue 2	djols djoft	shared root		
Blue 1 Purple 1	djols djonz	shared root	C versus D	Across category
Blue 1 Purple 1	sniks tsalp	different labels		
Blue 1 Purple 1	djols djonz	shared root	C versus C*	Perceptual check luminance sensitivity
Blue 1 Purple 2	djols djonz	shared root		
Blue 1 Purple 1	sniks tsalp	different labels	D versus D*	Perceptual check luminance sensitivity
Blue 1 Purple 2	sniks tsalp	different labels		
Blue 1 Blue 2	-	-	AB versus E	Perceptual check hue sensitivity
Orange 1 Purple 1	-	-		

Table 2

Comparison levels used in Experiment 1.

of ETN and HTN participants were similar over all experimental blocks, and 2) both ETN and HTN participants saw their results gradually improving.

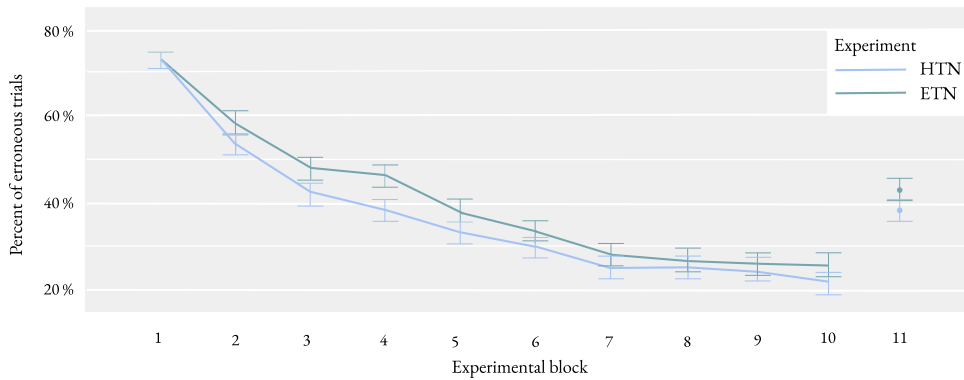


Figure 3: Average error rate over experimental blocks: Experiment 1
Session 1 (1st to 10th) and Session 2 (11th) of ETN and HTN.
Error bars represent the mean standard error.

Similarity ratings. Mean similarity ratings are reported in Figures 4, 5 and 6, and the statistical models are summarized in Tables 4, 5 and 6.

Overall model. We found a main effect of Group ($t = -3.275$) as participants rated the colour pairs as less similar between conditions in ETN compared to HTN group. We also found a main effect of Session number ($t = -7.718$), with colour pairs rated as being overall more similar during the second session compared to the first session. More importantly, within category separation was observed ($t = -2.463$) as participants rated nuances of the same hue that had received the same label as being more similar than when they had received two labels that only shared a root (e.g. Blue 1 and Blue 2 being either both *djols*, or *djonz*; *djoft*, respectively). Furthermore, we also observed between category compression ($t = -3.786$) since participants rated nuances belonging to different hues as more similar when their associated labels shared a root compared to when they were entirely different (e.g. Blue 1 and Purple 1 being *djonz*; *djoft* and *sniks*; *tsalp* respectively). Those results show that learned linguistic labels have a significant impact on perceived similarity among colour pairs (see Figure 4). The results also show an interaction between Group and categories of interest (Within category: $t = -2.496$; Across category: $t = -2.281$), showing that colour pairs were rated as more similar overall in HTN than in ETN. Finally, we saw an interaction between Group and Session number ($t = -4.185$), with colours rated as overall more similar in the second session compared to the first for ETN group.

The R formula is the following: $\text{Model} = \text{lmer}(\text{Rating}_c \sim \text{TrialNumber}_f + \text{Label}_f * \text{Group}_f * \text{SessionNumber}_f + (1|\text{ParticipantNumber}_f) + (1|\text{StimuliPresented}_f))$. Rating refers to similarity ratings; Trial Number accounts for the order of the stimuli displayed; Label refers to the label category; Group refers to ETN and HTN; Session Number

Language induced changes of mind

<i>Predictors</i>	<i>Estimates</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	0.5613	0.0869	0.4145 – 0.7602	-3.7317	<0.001
Experiment	0.9128	0.1412	0.6740 – 1.2361	-0.5898	0.555
Block 1 vs 2	0.4014	0.0586	0.3015 – 0.5345	-6.2492	<0.001
Blocks 1-2 vs 3	0.2771	0.0452	0.2012 – 0.3816	-7.8593	<0.001
Block 1-3 vs 4	0.3269	0.0563	0.2332 – 0.4582	-6.4921	<0.001
Blocks 1-4 vs 5	0.2119	0.0385	0.1484 – 0.3025	-8.5397	<0.001
Blocks 1-5 vs 6	0.1996	0.0377	0.1379 – 0.2890	-8.5339	<0.001
Blocks 1-6 vs 7	0.1435	0.0288	0.0969 – 0.2125	-9.6903	<0.001
Blocks 1-7 vs 8	0.2827	0.0411	0.2126 – 0.3760	-8.6860	<0.001
Blocks 1-8 vs 9	0.3645	0.0470	0.2832 – 0.4692	-7.8333	<0.001
Blocks 1-9 vs 10	0.3814	0.0453	0.3022 – 0.4813	-8.1212	<0.001
Session 1 vs Session 2	1.2085	0.1142	1.0042 – 1.4544	2.0047	0.045
Exp.*Blocks 1 vs 2	0.8845	0.1291	0.6644 – 1.1775	-0.8410	0.400
Exp.*Blocks 1-2 vs 3	0.9622	0.1570	0.6988 – 1.3249	-0.2359	0.813
Exp.*Block 1-3 vs 4	0.9134	0.1572	0.6518 – 1.2799	-0.5262	0.599
Exp.*Blocks 1-4 vs 5	1.0435	0.1894	0.7310 – 1.4894	0.2343	0.815
Exp.*Blocks 1-5 vs 6	1.1279	0.2128	0.7793 – 1.6326	0.6381	0.523
Exp.*Blocks 1-6 vs 7	1.1581	0.2318	0.7823 – 1.7143	0.7332	0.463
Exp.*Blocks 1-7 vs 8	1.2189	0.1771	0.9168 – 1.6206	1.3622	0.173
Exp.*Blocks 1-8 vs 9	1.1685	0.1505	0.9079 – 1.5040	1.2097	0.226
Exp.*Blocks 1-9 vs 10	1.0199	0.1210	0.8083 – 1.2869	0.1664	0.868
Exp.*Session 1 vs Session 2	0.9645	0.0911	0.8014 – 1.1607	-0.3826	0.702
Random Effects					
Residual variance	3.2899				
Intercept Stimuli Presented _f	0.0850				
Intercept Participant Number _f	1.1571				
ICC	0.2741				
^N Participant Number _f	66				
^N Stimuli Presented _f	16				
Observations	5808				
Marginal R ² / Conditional R ²	0.119 / 0.360				

Table 3

Model of naming accuracy: Experiment 1.

The R formula is the following: NA Model = *glmer(ErrorRate ~ TrialNumber + Experiment_f * BlockNumber_f + (1|ParticipantNumber_f) + (1|StimuliPresented_f), family = binomial)*.

defines the day of testing (1 or 2); Participant Number accounts for inter and intra participant variability; Stimuli Presented refers to the colour pair displayed.

Given the main effect of Group and the interaction between Group and the two variables of interest (Within and Across categories), we ran separate models for each group. The structure of the separate models remained the same.

ETN. We found a main effect of Within category ($t = -3.614$) and Across category ($t = -4.490$), respectively showing that 1) learning two labels for nuances of the same hue decreased the perceived similarity and 2) learning one unique label for two different hues increased the perceived similarity. We also found a main effect of Session number ($t = -8.883$) with participants rating the colour pairs as being overall more similar during the second session. The results show an interaction between Within category and Session number ($t = -2.146$), with participants rating the nuances of the same hue with two labels significantly more similar in the second session compared to the first session (see Figure 5).

The interaction between Across category and Session number was not significant ($t = -1.418$), showing that in the case of different colours having the same label, participants rated the colour pairs overall similarly between the first and second session.

HTN. We found no main effect of Within category ($t = 0.023$) and Across category ($t = -0.996$), with participants showing respectively no differences in their ratings if the nuances of a same colour had two labels or if two different colours were given a same label. However, a main effect of Session number was found ($t = -2.387$) with participants rating the colour pairs as being overall more similar during the second session. The interactions between Within

Language induced changes of mind

<i>Predictors</i>	<i>Estimates</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	0.9948	0.0854	0.8275 – 1.1622	11.6545	<0.001
Trial number	-0.0060	0.0198	-0.0448 – 0.0328	-0.3040	0.761
Exp.	-0.2797	0.0854	-0.4471 – -0.1123	-3.2750	0.001
Session	-0.1189	0.0154	-0.1491 – -0.0887	-7.7182	<0.001
Within category	-0.1597	0.0648	-0.2868 – -0.0326	-2.4633	0.014
Across category	-0.3328	0.0879	-0.5051 – -0.1605	-3.7863	<0.001
Percep. check lum. sensitivity	-0.7164	0.1117	-0.9353 – -0.4975	-6.4142	<0.001
Percep. check lum. sensitivity	-0.6071	0.1131	-0.8287 – -0.3855	-5.3693	<0.001
Percep. check hue sensitivity	4.3536	0.1197	4.1189 – 4.5883	36.3573	<0.001
Placeholder contrast	-0.8563	0.2658	-1.3773 – -0.3354	-3.2217	0.001
Exp.*Within category	-0.1619	0.0648	-0.2890 – -0.0348	-2.4963	0.013
Exp.*Across category	-0.2005	0.0879	-0.3728 – -0.0282	-2.2810	0.023
Exp.*Percep. check lum. sensitivity	0.1971	0.1117	-0.0218 – 0.4160	1.7649	0.078
Exp.* Percep. check lum. sensitivity	0.1655	0.1131	-0.0561 – 0.3872	1.4639	0.143
Exp.*Percep. check hue sensitivity	0.3889	0.1197	0.1542 – 0.6236	3.2476	0.001
Exp.*Placeholder contrast	-1.9923	0.2658	-2.5133 – -1.4714	-7.4956	<0.001
Session*Within category	-0.1005	0.0629	-0.2237 – 0.0228	-1.5982	0.110
Session*Across category	-0.1001	0.0861	-0.2689 – 0.0687	-1.1625	0.245
Session* Percep. check lum. sensitivity	0.0416	0.0611	-0.0782 – 0.1613	0.6803	0.496
Session* Percep. check lum. sensitivity	-0.0123	0.0607	-0.1313 – 0.1066	-0.2030	0.839
Session* Percep. check hue sensitivity	0.0694	0.0469	-0.1614 – 0.0226	-1.4790	0.139
Session*Placeholder contrast	-0.0788	0.1045	-0.2835 – 0.1260	-0.7542	0.451
Exp*Session	-0.0645	0.0154	-0.0947 – -0.0343	-4.1848	<0.001
Within category*Exp.*Session	-0.0854	0.0629	-0.2086 – 0.0379	-1.3577	0.175
Across category*Exp.*Session	-0.0609	0.0861	-0.2297 – 0.1079	-0.7074	0.479
Percep. check lum. sensitivity*Exp.*Session	0.0559	0.0611	-0.0639 – 0.1757	0.9152	0.360
Percep. check lum. sensitivity*Exp.*Session	-0.0125	0.0607	-0.1315 – 0.1064	-0.2066	0.836
Placeholder contrast*Exp.*Session	-0.2431	0.1045	-0.4479 – -0.0383	-2.3270	0.020
Percep. check hue sensitivity*Exp.*Session	-0.0334	0.0469	-0.1253 – 0.0586	-0.7110	0.477
Random Effects					
Residual variance	0.9140				
Intercept Stimuli Presented _f	0.0809				
Intercept Participant Number _f	0.3793				
ICC	0.3349				
NParticipant Number _f	66				
NStimuli Presented _f	105				
Observations	6582				
Marginal R ² / Conditional R ²	0.548 / 0.700				

Table 4

Overall model of similarity ratings: Experiment 1.

category and Session number, as well as Across category and Session number, were not significant (respectively $t = -0.167$ and $t = -0.305$), showing that participants did not rate the colours differently during the second session compared to the first session (see Figure 6).

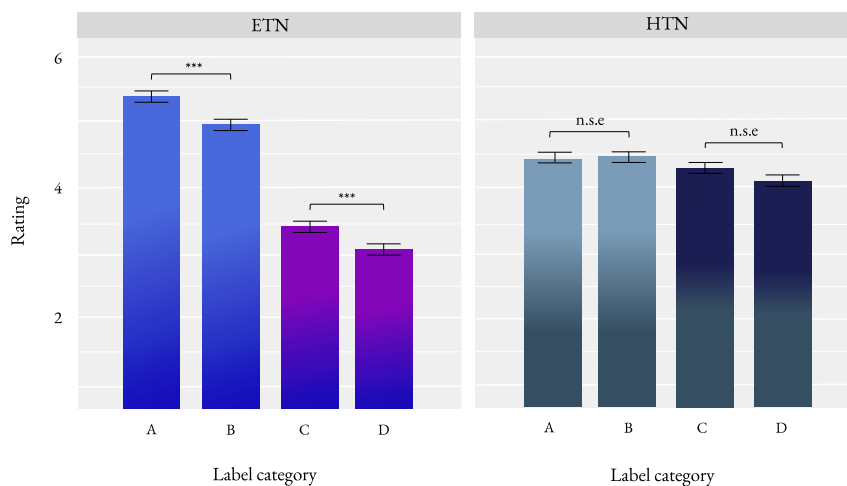


Figure 4: Similarity ratings of ETN and HTN for Experiment 1.

A vs B: within category separation; C vs D: between category compression.

Error bars represent the mean standard error. The colour scheme used in the following graphs is displayed as an example of pairs presented for each category (e.g., the first bar represents the within category compression, with both nuances of blue - Blue 1 at the bottom, Blue 2 at the apex - being given the same label (*djols*).

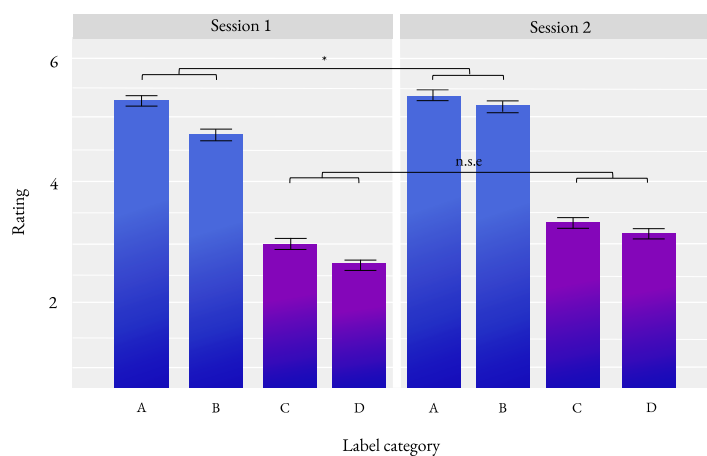


Figure 5: Similarity ratings over session 1 and 2 for ETN, Experiment 1.

A vs B: within category separation; C vs D: between category compression.

Error bars represent the mean standard error.

<i>Predictors</i>	<i>Estimates</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	0.7151	0.1027	0.5138 – 0.9165	6.9610	<0.001
Trial number	-0.0017	0.0274	-0.0555 – 0.0521	-0.0626	0.950
Session	-0.1835	0.0207	-0.2240 – -0.1431	-8.8834	<0.001
Within category	-0.3276	0.0907	-0.5053 – -0.1500	-3.6142	<0.001
Across category	-0.5300	0.1180	-0.7613 – -0.2986	-4.4898	<0.001
Percep. check lum. sensitivity	-0.4920	0.1252	-0.7374 – -0.2465	-3.9288	<0.001
Percep. check lum. sensitivity	-0.4076	0.1285	-0.6595 – -0.1557	-3.1713	0.002
Percep. check hue sensitivity	4.7425	0.1380	4.4721 – 5.0130	34.3727	<0.001
Placeholder contrast	-2.8519	0.3004	-3.4407 – -2.2630	-9.4924	<0.001
Session*Within category	-0.1856	0.0865	-0.3551 – -0.0161	-2.1462	0.032
Session*Across category	-0.1614	0.1138	-0.3846 – 0.0617	-1.4182	0.156
Session* Percep. check lum. sensitivity	0.0974	0.0806	-0.0606 – 0.2553	1.2083	0.227
Session* Percep. check lum. sensitivity	-0.0253	0.0804	-0.1828 – 0.1322	-0.3149	0.753
Session* Percep. check lum. sensitivity	-0.1030	0.0647	-0.2298 – 0.0238	-1.5926	0.111
Session*Placeholder contrast	-0.3221	0.1409	-0.5983 – -0.0458	-2.2850	0.022
Random Effects					
Residual variance	0.8242				
Intercept Stimuli Presented _f	0.0460				
Intercept Participant Number _f	0.2923				
ICC	0.2910				
^N Participant Number _f	34				
^N Stimuli Presented _f	49				
Observations	3163				
Marginal R ² / Conditional R ²	0.579 / 0.702				

Table 5
ETN model of similarity ratings: Experiment 1.

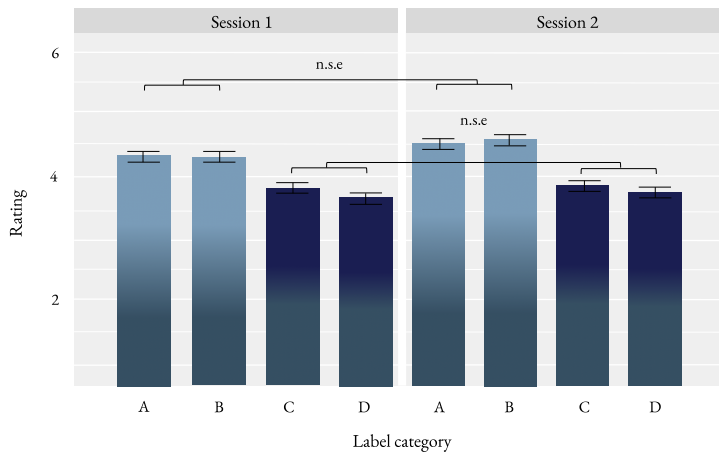


Figure 6: Similarity ratings over session 1 and 2 for HTN, Experiment 1.
A vs B: within category separation; C vs D: between category compression.
Error bars represent the mean standard error.

<i>Predictors</i>	<i>Estimates</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	1.2743	0.1370	1.0058 – 1.5427	9.3032	< 0.001
Trial number	-0.0028	0.0277	-0.0572 – 0.0516	-0.1013	0.919
Session	-0.0544	0.0228	-0.0990 – -0.0097	-2.3869	0.017
Within category	0.0021	0.0912	-0.1767 – 0.1810	0.0234	0.981
Across category	-0.1288	0.1294	-0.3824 – 0.1248	-0.9955	0.319
Percep. check lum. sensitivity	-0.9133	0.1909	-1.2875 – -0.5391	-4.7833	< 0.001
Percep. check lum. sensitivity	-0.7733	0.1905	-1.1466 – -0.4000	-4.0599	< 0.001
Percep. check hue sensitivity	3.9645	0.1888	3.5944 – 4.3346	20.9946	< 0.001
Placeholder contrast	1.1364	0.4281	0.2974 – 1.9754	2.6548	0.008
Session*Within category	-0.0151	0.0906	-0.1927 – 0.1625	-0.1670	0.867
Session*Across category	-0.0395	0.1292	-0.2927 – 0.2137	-0.3057	0.760
Session* Percep. check lum. sensitivity	-0.0147	0.0919	-0.1948 – 0.1654	-0.1597	0.873
Session* Percep. check lum. sensitivity	0.0000	0.0909	-0.1781 – 0.1781	0.0004	1.000
Session* Percep. check hue sensitivity	-0.0359	0.0674	-0.1681 – 0.0963	-0.5321	0.595
Session*Placeholder contrast	0.1643	0.1536	-0.1367 – 0.4654	1.0698	0.285
Random Effects					
Residual variance	0.9971				
Intercept Stimuli Presented _f	0.1120				
Intercept Participant Number _f	0.4714				
ICC	0.3691				
^N Participant Number _f	32				
^N Stimuli Presented _f	56				
Observations	3419				
Marginal R ² / Conditional R ²	0.514 / 0.693				

Table 6
HTN model of similarity ratings: Experiment 1.

2.4. Discussion

In Experiment 1, participants conducted a learning task followed by a similarity rating task to assess several aspects related to the impact of linguistic learning on categorical perception. Concretely, we asked whether the robustness of previously existing labels would modulate potential effects through co-activation and/or through mechanisms of bilingual language control triggered by such co-activation. Furthermore, we wanted to assess the impact of language learning on two different phenomena related to categorical perception, namely within category separation (i.e., creation of a new categorical boundary) and between category compression (i.e., blurring or elimination of a pre-existing categorical boundary). Finally, we also wanted to assess whether potential effects would last beyond the experimental session and without any immediately preceding task involving the colours and the labels. We observed that learning a new label for colours with a previous robust lexical label (ETN) resulted in both within category separation and between category compression effects. No such effects were present when learning new labels for colours without a robust previous label (HTN). The effects observed for the ETN colours were larger on the first day in the case of within category separation, but not different across days in the case of between category compression. Importantly, though, both types of effects were present across the two sessions. Thus, we show that the impact of language learning on categorical perception is very pervasive, affecting categorical boundaries in both directions.

Still, with the current experimental design, we could not directly compare the magnitude of the language induced effects within the same category and across different categories. Descriptively, it looks as if the effects of within category separation are larger than those of between category compression, but as the manipulations of these two contrasts were not comparable it is impossible to say for certain. We will address this comparison of within category separation and between category suppression directly in Experiment 2. We also show that the effect of language on categorical perception is rather long-lasting and not dependent upon the immediately preceding experience. This suggests that the phenomenon observed here is related to authentic learning and not only to transient stimulus response mappings.

Concerning the interaction of the effects with lexical robustness of previous labels, one might wonder whether this may be due to the HTN colours being overall harder to identify than the ETN colours, thus making them harder to perceive and associate with a label. However, two aspects of the data can be used to argue against this possibility. First, even though the HTN colour group did not show any effect of the newly learnt labels on categorical perception, participants did show sensitivity to other aspects of perceptual variance such as differences in luminance and differences in hue (see Figure ?? in Additional Material). Finally, even though learning in the ETN colour group was descriptively slightly better, this difference was not statistically significant. Note though that we had very few observations to assess naming accuracy (one per item and block for each participant), so it is possible that with more observations we would have observed a difference in naming accuracy between the group. Nevertheless, participants in the HTN group attained the same level of learning accuracy as those of the ETN group by the end of the learning part of the experiment, and there was no difference in retention rate across groups. Hence, insufficient learning or perception of the HTN labels and colours respectively cannot account for the observed effects. The results can be accounted for by the hypothesis assuming that the lexico-semantic links of new labels are strengthened in proportion to the co-activation of previous labels. This would result in a much larger strengthening in the case of ETN colour labels compared to HTN colour labels. However, according to this hypothesis, one would expect better learning of the ETN labels as a consequence of greater strengthening. This hypothesis will be further examined in Experiment 2 by increasing the number of observations in the testing phase of the naming out loud task.

3. Experiment 2

To verify the robustness of the effects of linguistic learning on categorical perception observed in Experiment 1, we conducted a second experiment. By slightly modifying the experimental design, the aim of Experiment 2 was to maximize differences in categorical perception driven by linguistic learning. Contrary to Experiment 1, the Root sharing category was suppressed: e.g., two nuances of blue would either share a same label (e.g., *djols*) or have distinct labels (e.g., *sniks*; *tsalp*). The same label manipulation was applied to colour pairs belonging to the same hue (e.g., Blue 1 and Blue 2) and colour pairs belonging to different hues (e.g., Blue 1 and Purple 1). We predicted that colours receiving two labels would be rated as perceptually less similar than colours receiving a same label. Furthermore, thanks to our modified 2x2 design, we were able to directly test the interaction between Label category and Colour category. This would shed light on whether the language induced effects were larger on within category separation or on between category compression. Finally, we also introduced some changes to the naming out loud (NOL) part of the experiment. First, to be able to reliably compare naming accuracy performance across the ETN and HTN groups,

we increased the number of testing repetitions per block from 1 to 3. Second, in order to keep the difficulty across Experiments 1 and 2 comparable and given that in this new experiment the participant only had to learn 4 labels (instead of 8 as in Experiment 1), we reduced the number of learning repetitions from 5 to 3, and the number of experimental blocks from 10 to 6.

3.1. Material and methods

Participants. Due to an experiment spamming, 18 participants were removed from analysis, resulting in testing 18 additional participants. In total, 90 participants giving written consent and participating in the experiment ($M = 25.43$ years; $SD = 3.35$; [18; 31]; 64 women, 26 men). Along with the 18 participants that spammed, two participants were rejected due to equipment failure (microphone issues), leaving a sample of 70 participants for analysis (35 for ETN, 35 for HTN). All participants were native French-speakers recruited through Aix-Marseille University and were compensated 10 euros via PayPal through the FindingFive platform. The experiment took an average of 40 minutes to be completed over one single session. All participants reported no hearing, speech or neurological disorders, and had normal or corrected-to-normal vision, with no colour perception disorder. Two participants were rejected due to equipment failure (microphone issues), leaving a sample of 70 participants for analysis (35 for ETN, 35 for HTN). As in Experiment 1, participants were asked to self-assess their language proficiency through a preliminary questionnaire (see details in Additional Material).

Experimental design. The experiment involved a 2×2 design with one independent variable being the label category (two levels: one label, two labels) and the other independent variable being the colour category (two levels: across, within). As in Experiment 1, the robustness of previous linguistic labels was manipulated between participants by using easy to name (ETN) and hard to name (HTN) colours, respectively. The dependent variables were naming accuracy (for the learning phase) and the similarity ratings.

Stimuli. The same 16 colour stimuli were used as in Experiment 1, but the label set was different. Only four labels were used, three from the Experiment 1 (*sniks*, *djols*, *tsalp*) plus a new one (*flurn*). The newly added label was tested using the same procedure as in Experiment 1.

3.2. Instrumentation and measures

The same instrumentation was used as in Experiment 1.

Procedures. Participants began the experiment with a label-colour association (naming out loud task), followed by a similarity rating task.

Naming out loud task. Each participant completed six blocks of the naming out loud task. The order of stimuli within the blocks was randomized across and within participants. The naming out loud task was divided into a learning phase and a testing phase. In the learning phase, the eight label-colour associations were randomly presented three times each, and displayed at the centre of the screen in a randomized order in the six experimental blocks, so a total of 180 learning trials. For each colour, the label colour association was presented for 1000 ms, followed by the colour chip for 1500 ms, during which the participant was instructed to name out loud the label associated with the colour displayed. A blank inter-trial interval of 700 ms intervened between each trial. At the end of each learning phase, a testing phase was presented, in which all eight colour chips were displayed three times each for 3000 ms, during which the participant was instructed to name out loud the label associated with the colour displayed. The testing trials were composed of a total of 144 stimuli, with three randomized presentations of each stimulus at the end of every experimental block. This task took a total of 22 minutes.

Similarity rating task. The similarity rating procedure was the same as in Experiment 1.

Statistical analysis. The goal of the current study was to 1) replicate the findings of Experiment 1 and 2) create a reduced 2×2 design to directly address the interaction between Label and Colour categories. All analyses were the same as Experiment 1 except for the condition Label, which was specially coded to address the comparisons of interest, i.e., A = same label same hue, B = different label same hue, C = same label different hue, D = different label different hue. The crucial contrasts to assess our hypotheses are the main effects of colour category (A-B vs C-D) label category (A-C vs B-D), as well as their interaction. See Table 7 for comparison levels. Within category separation and between category compression were assessed through post-hoc tests (see section 3.3 below).

3.3. Results

Naming accuracy. Mean naming accuracy is reported in Figure 7 and the full results in Table 8. The results show a main effect of Group ($t = -3.8173$), with ETN participants performing overall better than HTN participants.

Language induced changes of mind

Colour pairs	Labels	Label category	Comparison levels	Category
Blue 1 Blue 2	djols djols	same label	A versus B	Within category
Blue 1 Blue 2	sniks tsalp	different labels		
Blue 1 Purple 1	djols djols	same label	C versus D	Across category
Blue 1 Purple 1	sniks tsalp	different labels		
Blue 1 Blue 2	-	-	AB versus CD	Colour category
Blue 1 Purple 1	-	-		
Blue 1 Blue 1	-	-	AC versus BD	Label category
Purple 1 Purple 1	-	-		

Table 7

Comparison levels used in Experiment 2.

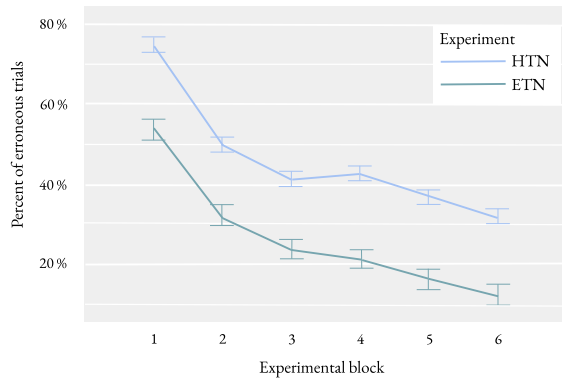


Figure 7: Average error rate over experimental blocks (1st to 6th) of ETN and HTN for Experiment 2. Error bars represent the mean standard error.

No interaction between Block number and Group was found, but a main effect of Block number was present for each of the six experimental blocks, showing that participants in both groups saw their performance gradually improving.

Similarity ratings. Mean similarity ratings are reported in Figure 8, and the statistical models are summarized in Tables 9, 10 and 11.

Overall model. We found a main effect of Label category ($t = 3.365$) as participants rated the colour pairs which were given one label as more similar than colour pairs with two labels. We also found a main effect of Colour category ($t = -3.668$), with participants rating the nuances of a same colour as more similar than colours belonging to distinct categories. We found a main effect of Group ($t = -2.0591$), showing that overall, significant differences in the ratings were observable between ETN and HTN groups. The results also show an interaction between 1) Group and Label category ($t = 2.871$) and 2) Group and Colour category ($t = -3.69$), showing respectively that the ratings of colours categories and label categories differed between ETN and HTN (see below for additional models). Finally, we found no interaction between Label category and Colour category ($t = 1.6599$), showing that nuances of a same colour being given a single label were not rated as more similar than two colours being given two different labels. Given the interaction between Group and the two variables of interest (colour and label categories), we decided to run separate models for each group.

The R formula is the following: $\text{Model} = \text{lmer}(\text{Rating}_c \sim \text{Trial Number} + \text{Label Category}_f * \text{Colour Category}_f * \text{Group}_f + (1|\text{Participant Number}_f) + (1|\text{Stimuli Presented}_f))$.

Rating refers to similarity ratings; Trial Number accounts for the order of the stimuli displayed; Label category refers to the label (same vs different); Colour category refers to the colour (nuances of same colour vs different colours); Group refers to ETN and HTN; Participant Number accounts for inter and intra participant variability; Stimuli Presented refers to the colour pair displayed.

ETN. We found a main effect of Label category ($t = 3.984$) and Colour category ($t = -6.322$), showing respectively that 1) learning two labels decreased the perceived similarity between colours compared to learning a single label, and

Language induced changes of mind

<i>Predictors</i>	<i>Pdd ratios</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	0.4843	0.1361	0.2793 – 0.8400	-2.5807	0.010
Experiment	0.6674	0.1065	0.4882 – 0.9124	-2.5348	0.011
Block 1 vs 2	0.2774	0.0270	0.2293 – 0.3356	-13.1927	<0.001
Blocks 1-2 vs 3	0.2170	0.0243	0.1742 – 0.2702	-13.6412	<0.001
Block 1-3 vs 4	0.2337	0.0283	0.1843 – 0.2964	-11.9941	<0.001
Blocks 1-4 vs 5	0.1899	0.0247	0.1471 – 0.2450	-12.7650	<0.001
Blocks 1-5 vs 6	0.1288	0.0184	0.0974 – 0.1704	-14.3565	<0.001
Exp.*Blocks 1 vs 2	1.1759	0.1139	0.9726 – 1.4218	1.6729	0.094
Exp.*Blocks 1-2 vs 3	1.1347	0.1266	0.9118 – 1.4120	1.1324	0.257
Exp.*Block 1-3 vs 4	0.9042	0.1092	0.7136 – 1.1458	-0.8332	0.405
Exp.*Blocks 1-4 vs 5	0.9139	0.1185	0.7088 – 1.1783	-0.6946	0.487
Exp.*Blocks 1-5 vs 6	0.8891	0.1263	0.6730 – 1.1745	-0.8278	0.408
Random Effects					
Residual variance	3.2899				
Intercept Target Expected _f	0.3559				
Intercept Participant Number _f	1.2724				
ICC	0.3311				
^N Participant Number _f	53				
^N Target Expected _f	7				
Observations	7632				
Marginal R ² / Conditional R ²	0.155 / 0.435				

Table 8

Model of naming accuracy: Experiment 2.

The R formula is the following: NA Model = *glmer(ErrorRate ~ TrialNumber + Experiment_f * BlockNumber_f + (1|ParticipantNumber_f) + (1|StimuliPresented_f), family = binomial)*.

<i>Predictors</i>	<i>Estimates</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	-0.0179	0.1705	-0.3520 – 0.3163	-0.1047	0.917
Trial number	-0.0090	0.0431	-0.0935 – 0.0755	-0.2079	0.835
Exp.	-0.2723	0.1705	-0.6064 – 0.0618	-1.5973	0.110
Label category	0.1506	0.0450	0.0624 – 0.2388	3.3464	0.001
Colour category	-0.2825	0.0900	-0.4588 – -0.1061	-3.1400	0.002
Exp.*Label category	0.1420	0.0450	0.0538 – 0.2302	3.1563	0.002
Exp.*Colour category	-0.2911	0.0899	-0.4674 – -0.1148	-3.2361	0.001
Label category*Colour category	0.0962	0.0450	0.0081 – 0.1844	2.1394	0.032
Exp.*Label category*Colour category	0.0505	0.0450	-0.0377 – 0.1387	1.1231	0.261
Random Effects					
Residual variance	1.4594				
Intercept Stimuli Presented _f	0.2021				
Intercept Participant Number _f	1.0921				
ICC	0.4700				
^N Participant Number _f	53				
^N Stimuli Presented _f	32				
Observations	848				
Marginal R ² / Conditional R ²	0.096 / 0.521				

Table 9

Overall model of similarity ratings: Experiment 2.

Assessing direct interaction between Label and Colour categories.

2) colours belonging to distinct categories were rated as more dissimilar than nuances of a same colour (e.g., Blue 1 and Purple 1 compared to Blue 1 and Blue 2). We saw no interaction between Label and Colour categories ($t = 0.9093$) showing that the ratings between colour pairs belonging to the same hue (e.g., Blue 1 and Blue 2) were overall not more similar than the ratings of colours of different categories (e.g., Blue 1 and Purple 1), in the case of the colours having one or two labels.

Language induced changes of mind

Predictors	Estimates	Std err	CI	t-value	p
Intercept	3.6951	0.2218	3.2604 – 4.1297	16.6626	<0.001
Trial number	-0.1012	0.0727	-0.2437 – 0.0412	-1.3931	0.164
Label category	0.2858	0.0755	0.1378 – 0.4339	3.7845	<0.001
Colour category	-0.5776	0.1105	-0.7942 – -0.3610	-5.2264	<0.001
Label category*Colour category	0.1455	0.0754	-0.0024 – 0.2933	1.9287	< 0.054
Random Effects					
Residual variance	1.8113				
Intercept Stimuli Presented _f	0.1152				
Intercept Participant Number _f	0.8501				
ICC	0.3699				
^N Participant Number _f	23				
^N Stimuli Presented _f	16				
Observations	368				
Marginal R ² / Conditional R ²	0.148 / 0.444				

Table 10
ETN model of similarity ratings: Experiment 2.

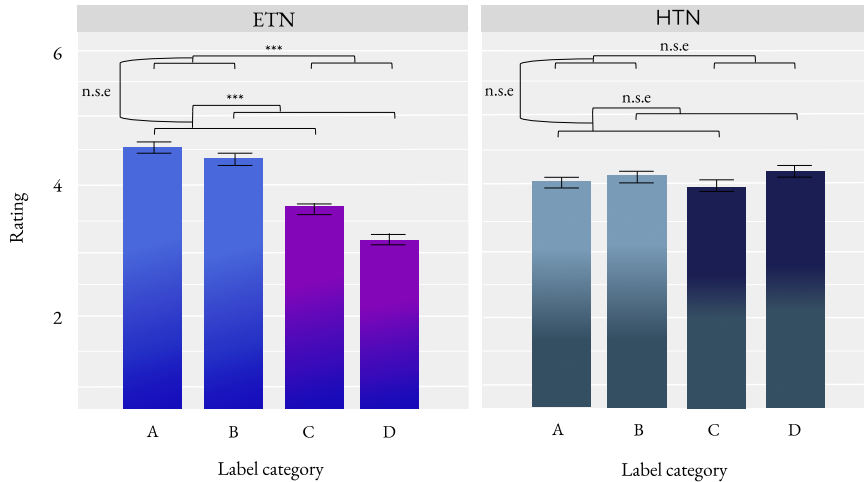


Figure 8: Similarity ratings of ETN and HTN for Experiment 2.

AC vs BD: label category; AB vs CD: colour category. AB-CD vs AC-BD: interaction between label and colour categories. Error bars represent the mean standard error.

HTN. We found no main effect of Label category ($t = 0.532$) or Colour category ($t = -0.024$), as well as no interaction between colour and label categories ($t = 1.637$), showing that nuances were not rated differently whether they were given the same or different labels, and two colours were not rated differently depending on whether they received the same or different labels.

The R formula is the following:

Model = $\text{Ime}(\text{Rating}_c \sim \text{TrialNumber} + \text{Label}_f * \text{Group}_f + (1|\text{ParticipantNumber}_f) + (1|\text{StimuliPresented}_f)$. Rating refers to similarity ratings; Trial Number accounts for the order of the stimuli displayed; Label refers to the label category (same vs different); Group refers to ETN and HTN; Participant Number accounts for inter and intra participant variability; Stimuli Presented refers to the colour pair displayed.

Post-hoc tests. Given the ETN results, post-hoc tests assessing the Within and Across categories effect were conducted. Mean similarity ratings are reported in Figure 9.

We found an effect of Within category ($t = 2.37$), showing that nuances of the same colour being given one label (*djols*) were rated more similarly than nuances with two labels (*sniks; tsalp*). We also found an effect of Across category

Language induced changes of mind

<i>Predictors</i>	<i>Estimates</i>	<i>Std err</i>	<i>CI</i>	<i>t-value</i>	<i>p</i>
Intercept	4.2403	0.2486	3.7530 – 4.7277	17.0543	<0.001
Trial number	0.0585	0.0517	-0.0428 – 0.1599	1.1320	0.258
Label category	0.0190	0.0523	-0.0835 – 0.1214	0.3628	0.717
Colour category	0.0122	0.1408	-0.2638 – 0.2883	0.0870	0.931
Label category*Colour category	0.0552	0.0523	-0.0473 – 0.1576	1.0551	0.291
Random Effects					
Residual variance	1.1866				
Intercept Stimuli Presented _f	0.2775				
Intercept Participant Number _f	1.2596				
ICC	0.5644				
^N Participant Number _f	30				
^N Stimuli Presented _f	16				
Observations	480				
Marginal R ² / Conditional R ²	0.002 / 0.565				

Table 11
HTN model of similarity ratings: Experiment 2.

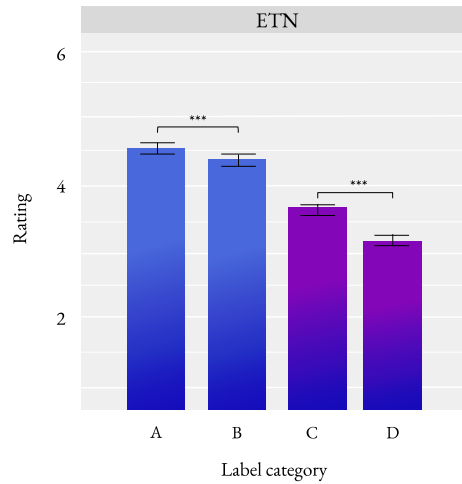


Figure 9: Similarity ratings of ETN for Experiment 2, post-hoc tests.
A vs B: within category separation; C vs D: between category compression.
Error bars represent the mean standard error.

was found ($t = 3.899$), showing that colours belonging to distinct categories having the same label (*djols*) were rated overall more similar than colours with different labels (*sniks; tsalp*). Comparison levels and model outputs are available in the Additional Material.

3.4. Discussion

In Experiment 2, we again found that the linguistic similarity (i.e., same or different label) with which two colours were named had an impact on categorical perception: participants rated colour pairs as being more similar if they were given the same label compared to two different labels. As in Experiment 1, this effect was only present for the colours with a previous robust label (ETN).

Interestingly, the label manipulation did not interact with the colour category manipulation. A follow-up analysis showed that the label effect on categorical perception was present in the Across condition, where the colours belonged to different categories (e.g., blue and purple), as well as the Within condition, where the colours belonged to the same category (e.g., light blue and dark blue). The fact that the label effect was similar for both Within and Across

suggests that language induced effects on categorical perception are effective both in blurring prior category boundaries (between category compression) and in establishing new ones (within category separation).

Finally, though participants of both the ETN and HTN groups showed significant learning throughout the naming out loud session, participants were significantly more accurate for the ETN compared to the HTN along all experimental blocks. Just as with the larger effect of the label manipulation in the ETN group, this shows that the previous native label attached to the colour did not hinder the learning of new labels. This lends further support to the hypothesis that during vocabulary learning, the lexico-semantic links of new words are strengthened in proportion to the co-activation of previous words attached to the object being named.

4. General Discussion

The aim of this study was to increase our understanding of the nature and the boundaries of effects on the categorical perception of colours induced by a short language learning experience. In two experiments, participants learned new labels for two nuances of four hues. Across participants, we manipulated whether the colours had a previous robust label (ETN colours) or not (HTN colours). The linguistic similarity of labels for different nuances and hues was manipulated within participants. Upon completion of learning, participants rated the similarity of the colours they just learned to label presented in pairs. In Experiment 1, this similarity rating was repeated after 24 hours. In a nutshell, we found that colours that had a previous robust label were persistently rated as more similar if they had similar new labels, and as more dissimilar if they had dissimilar new labels. In what follows, we will discuss how our results contributed to refine our knowledge of second language learning and reorganization of perceptual categories.

A first contribution of this study relates to whether effects on categorical perception would be observed when learning a new label for an object that already has a label in the native language. Our results showed that learning new labels for colours with a robust previous lexical label (ETN colours) had a consistent impact on categorical perception, while no such effect was observed for colours with a less robust previous lexical label (HTN colours). In the introduction, we had put forward three hypotheses resulting from properties inherent to the dynamics of (bilingual) language production. In fact, whenever a speaker wants to verbalize a concept, related concepts and their corresponding words also become active through spreading activation (e.g., Roelofs). Such spreading activation is thought to be especially challenging in the case of bilingual speakers or language learners for whom the words of the native language are much more readily accessible and are thought to become co-activated even in contexts in which only the other language is used. As discussed in the Introduction, there are potentially three ways in which speakers might deal with the co-activation of previous words when learning a new language. The first possibility is that speakers rely on the previous labels in order to learn the new ones through translation, as proposed by the revised hierarchical model (e.g., Kroll and Stewart (1994)). However, in our case this strategy would likely not have been very effective because there was only a correspondence between the organization of the semantic space of participants' native language and the new vocabulary for half of the items. Furthermore, as this account involves keeping the previous lexico-semantic mappings active, presumably the impact of new lexico-semantic mappings on categorical perception would likely be reduced or even absent. A second possibility is by inhibiting the co-active non-target words in proportion to their activation (e.g., Green (1998) for an inhibitory model of bilingual language control; but see also F. for a monolingual model with weakening of the lexico-semantic links of semantic competitors). This explanation, if anything, leads to the prediction that words with strong competitors requiring much inhibition would be harder to learn and potentially lead to smaller language-induced effects on categorical perception. A third possibility and the only one that can account for our findings is that the lexico-semantic links of the new lexical items are strengthened in proportion to the co-activation of previous lexical items (e.g., Oppenheim Dell, 2010; Runnqvist et al. (2019)). It will be interesting to further examine differences in learning and interactions with categorical perception as a function of the robustness of previous labels in future research to assess whether with more practice the effects would be similar.

A second contribution of this study relates to the question of whether the effects of learning a new language on categorical perception extend to within category separation and between category compression. Overall, in the ETN experiments we found that colours that have been given two labels were rated significantly more dissimilar than the colours being given one label, both for nuances of a same colour and colours belonging to distinct categories. Interestingly, we were able to show that, everything else being equal, the effect of language on blurring categorical boundaries (between category compression) is larger than the effect of creating a new boundary within a category (within category separation). This is robust evidence that categorical perception is affected by language, as the between category compression dissociates perceptual similarity and linguistic similarity.

A third contribution of this study relates to the question of whether language-induced effects on categorical perception obtained in the context of a single experimental session reflect actual persistent lexical learning and reorganization of perceptual categories or only a transient stimulus response association. The fact that the language learning impact on similarity ratings was still present after 24 hours and without any immediately preceding naming of the newly learnt labels indicate that the effects observed here go beyond any form of transient priming and rather had a persistent impact on the organization of the perceptual categories themselves. Given that these effects were obtained after a very short learning session (roughly half an hour) the most striking conclusion is that our perception of reality is permanently under construction and very easily influenced by external and arbitrary associations.

4.1. Limitations and future directions

Before concluding, some limitations should be noted. First, running online experiments on behavioural tasks, even though proven to be statistically valid (see [Fairs and Strijkers \(2021\)](#) on speech production), can induce some issues: Online experiments induce more variability in stimulus display, as the colours may have differed across the computer screens. Yet, even though frequently pointed out as being unsound because they are conducted with non-calibrated screens, studies like the one conducted by [Paramei, Griber and Mylonas \(2018\)](#) on colour naming have shown that online setups are comparable to offline experiments in terms of data quality. Third, it is possible that the absence of effects in the HTN group was due in part to the increased difficulty to discriminate the colours and thus successfully associate a label to them. Little research has been conducted on HTN colour perception so far, leaving unclear the implications of purely perceptual constraints on lexical categorization: often being muted tones, it is likely that HTN colours do not possess a robust label because they are harder to consistently name due to their lack of striking features. Consistent with this hypothesis, previous work by [Luke](#) on perceptual decision-making has showed that HTN colours are more difficult to map over labels. However, in our study, participants of both Experiment 1 and Experiment 2 achieved rather high naming accuracy for the HTN colours. Moreover, the perceptual checks of Experiment 1 seemed to preclude the possibility of perceptual discrimination difficulties to some extent as they showed that participants did perceive within-hue differences of luminance (light and dark grey-blue) and differences between hues that were far apart on the colour wheel (i.e., grey-blue and dust-gold). Still, only in Experiment 1 but not in Experiment 2 did they differentiate between closely related hues (i.e., grey-blue and lilac being perceived as less similar than light and dark grey-blue), suggesting that at least under certain circumstances these hues were less distinctly perceived to begin with. One possibility is of course that this is a type of language-induced effect on perception: in Experiment 2 (and not in Experiment 1), half of the closely related hues received the same label just as half of the nuances of the same hue, hence favouring similarity ratings to be comparable for these items. It is possible that this context favoured the observed similarity confusion. Further experimentation testing similarity ratings between pairs of colours prior to a linguistic learning phase would potentially help to remove ambiguity on perceptual and linguistic interferences. In addition, the results of the ongoing EEG study will be informative regarding pre-attentive perception sensitivity. In any event, the high naming accuracy is difficult to reconcile with perceptual difficulties, as is the fact that in Experiment 1 participants did rate nuances as more similar than close hues. Fourth, it is still to be discovered if the effects observed here can be extended to other aspects of human perception. Further research could focus on different stimuli, such as abstract concepts (e.g. space and time) or new semantic categories (e.g. inducing a perceptual shift in the semantic categorization of insects and four-legged mammals). Another line of research could address perceptual categories such as voice perception or, following the work of [Yu, Li, Mo and Mo \(2017\)](#), address the pre-attentive cues to face categorization.

In conclusion, the current study helped to gain insight on several aspects related to the reorganization of perceptual categories induced by language learning. Persistent and unambiguously language-induced effects on categorical perception of colours were obtained after a short learning session, but only for colours that had a previous robust label. These findings highlight 1) a key role of spreading activation and lexico-semantic strengthening induced by competition as a motor of second language learning, and 2) an impressive malleability of the human, perception-based interpretation of the world.

4.2. Funding

This work, carried out within the Institute of Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government (France 2030), managed by the French National Agency for Research (ANR) and

the Excellence Initiative of Aix-Marseille University (A*MIDEX). E.R. has benefited from support from the French government, managed by the French National Agency for Research (ANR) through a research grant (ANR-18-CE28-0013).

Bibliography

References

- Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, R., Thierry, G., 2010. Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition* 116, 437–443. doi:10.1016/j.cognition.2010.05.016.
- Baayen, R.H., Davidson, D.J., Bates, D.M., 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59, 390–412. doi:10.1016/j.jml.2007.12.005. place: Netherlands Publisher: Elsevier Science.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1–48. doi:10.18637/jss.v067.i01. number: 1.
- Braem, S., Verguts, T., Notebaert, W., 2011. Conflict adaptation by means of associative learning. *Journal of Experimental Psychology: Human Perception and Performance* 37, 1662–1666. doi:10.1037/a0024385. place: US Publisher: American Psychological Association.
- Colomé, J., 2001. Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language* 45, 721–736. doi:10.1006/jmla.2001.2793. place: Netherlands Publisher: Elsevier Science.
- Davis, M.H., Gaskell, M.G., 2009. A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364, 3773–3800. doi:10.1098/rstb.2009.0111.
- Dell, F., 1995. Consonant clusters and phonological syllables in French. *Lingua* 95, 5–26. doi:10.1016/0024-3841(95)90099-3.
- Duyck, W., Desmet, T., Verbeke, L.P.C., Brysbaert, M., 2004. WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers* 36, 488–499. doi:10.3758/BF03195595.
- F., G.M.M., . The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production 114, 227–252. doi:10.1016/j.cognition.2009.09.007.
- Fairs, A., Strijkers, K., 2021. Can we use the internet to study speech production? yes we can! evidence contrasting online versus laboratory naming latencies and errors .
- Gervits, F., Johanson, M., Papafragou, A., 2016. Intentionality and the Role of Labels in Categorization. *CogSci* .
- Green, D.W., 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1, 67–81. doi:10.1017/S1366728998000133. publisher: Cambridge University Press.
- Incera, S., McLennan, C.T., 2018. The time course of within and between-language interference in bilinguals. *International Journal of Bilingualism* 22, 88–99. doi:10.1177/1367006916644688. publisher: SAGE Publications Ltd.
- Jaeger, T.F., 2008. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of memory and language* 59, 434–446. doi:10.1016/j.jml.2007.11.007.
- Johanson, M., Papafragou, A., 2016. The influence of labels and facts on children's and adults' categorization. *Journal of Experimental Child Psychology* 144, 130–151. doi:10.1016/j.jecp.2015.11.010. publisher: Academic Press Inc.
- Johnson, R.L., Eisler, M.E., 2012. The importance of the first and last letter in words during sentence reading. *Acta Psychologica* 141, 336–351. doi:10.1016/j.actpsy.2012.09.013. place: Netherlands Publisher: Elsevier Science.
- Juárez, F.P.G., Sicotte, T., Thériault, C., Harnad, S., 2019. Category learning can alter perception and its neural correlates. *PLoS ONE* 14. doi:10.1371/journal.pone.0226000. publisher: Public Library of Science.
- Kroll, J.F., Bobb, S.C., Misra, M., Guo, T., 2008. Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychologica* 128, 416–430. doi:10.1016/j.actpsy.2008.02.001.
- Kroll, J.F., Stewart, E., 1994. Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language* 33, 149–174. doi:10.1006/jmla.1994.1008. place: Netherlands Publisher: Elsevier Science.
- Lima, S.D., Inhoff, A.W., 1985. Lexical access during eye fixations in reading: Effects of word-initial letter sequence. *Journal of Experimental Psychology: Human Perception and Performance* 11, 272–285. doi:10.1037/0096-1523.11.3.272. place: US Publisher: American Psychological Association.
- Luke, T.C.K.Y., . Language affects patterns of brain activation associated with perceptual decision 105, 4004–4009. doi:10.1073/pnas.0800055105.
- Lupyan, G., Abdel Rahman, R., Boroditsky, L., Clark, A., 2020. Effects of Language on Visual Perception. *Trends in Cognitive Sciences* 24, 930–944. doi:10.1016/j.tics.2020.08.005. publisher: Elsevier Ltd.
- Paramei, G.V., Griber, Y.A., Mylonas, D., 2018. An online color naming experiment in Russian using Munsell color samples. *Color Research and Application* 43, 358–374. doi:10.1002/co1.22190. publisher: John Wiley and Sons Inc.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Raz, A., Kirsch, I., Pollard, J., Nitkin-Kaner, Y., 2006. Suggestion Reduces the Stroop Effect. *Psychological Science* 17, 91–95. doi:10.1111/j.1467-9280.2006.01669.x. publisher: SAGE Publications Inc.
- Roelofs, J., A spreading-activation theory of lemma retrieval in speaking 42, 107–142. doi:10.1016/0010-0277(92)90041-f.
- Rolstad, K., MacSwan, J., 2014. The facilitation effect and language thresholds. *Frontiers in Psychology* 0. doi:10.3389/fpsyg.2014.01197. publisher: Frontiers.
- Runnqvist, E., Costa, A., 2012. Is retrieval-induced forgetting behind the bilingual disadvantage in speech production? *Bilingualism: Language and Cognition* 15, 365–377. URL: <https://hal.archives-ouvertes.fr/hal-01489703>, doi:10.1017/S1366728911000034.

- Runnqvist, E., Strijkers, K., Alario, F.X., Costa, A., 2012. Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language* 66, 850–869. doi:<https://doi.org/10.1016/j.jml.2012.02.007>.
- Runnqvist, E., Strijkers, K., Costa, A., 2019. Error-based learning and lexical competition in word production: Evidence from multilingual naming. *PLoS ONE* 14. doi:10.1371/journal.pone.0213765.
- Schmid, M.S., 2013. First language attrition. *WIREs Cognitive Science* 4, 117–123. doi:10.1002/wcs.1218.
- Stefanics, G., Astikainen, P., Czigler, I., 2015. Visual mismatch negativity (vMMN): A prediction error signal in the visual modality. *Frontiers in Human Neuroscience* 8. doi:10.3389/fnhum.2014.01074. publisher: Frontiers Media S. A.
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J., 2009. Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences* 106, 4567–4570. doi:10.1073/pnas.0811155106.
- Winawer, J., Witthoft, N., Frank, M.C., Wu, L., Wade, A.R., Boroditsky, L., 2007. Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences* 104, 7780–7785. doi:10.1073/pnas.0701644104.
- Wu, Y.J., Thierry, G., 2010. Chinese-English Bilinguals Reading English Hear Chinese. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30, 7646–51. doi:10.1523/JNEUROSCI.1602-10.2010.
- Yu, M., Li, Y., Mo, C., Mo, L., 2017. Newly learned categories induce pre-attentive categorical perception of faces. *Scientific Reports* 7. doi:10.1038/s41598-017-14104-6. publisher: Nature Publishing Group.