Longitudinal evidence for simultaneous bilingual language development with shifting language dominance, and how to explain it

Gary M. Oppenheim[*†] Zenzi M. Griffin[‡]
Elizabeth D. Peña[§] Lisa M. Bedore[¶]

in press at *Language Learning*

Abstract

Theories of how language works have shifted from rule-like competence accounts to more skill-like incremental learning accounts. Under these, people acquire language incrementally, through practice, and may even lose it incrementally as they acquire competing mappings. Incremental learning implies that (1) a bilingual's abilities in their languages should depend on how much they practice each (not merely age of acquisition), and (2) using an L2 more could cause a bilingual to gradually 'unlearn' their L1. Using timed picture naming and vocabulary measures, we tracked 139 children for several years as they transitioned from mostly-Spanish homes to mostly-English schools. Following their increased English use, many became more proficient in English than Spanish around the third grade, demonstrating continual learning. But their Spanish also improved, showing that L1-attrition is not inevitable. Incremental learning explains both co-improvement and L1-attrition as consequences of experience-driven learning: improvement from continuing L1 use can offset competitive unlearning.

[*]School of Psychology, Bangor University
[†]Department of Psychology, Rice University
[‡]Department of Psychology University of Texas at Austin
[§]School of Education University of California, Irvine
[¶]Department of Communication Sciences and Disorders, Temple University

Thanks to Tamar Gollan, J. Greg Hixon, members of the Bangor University Centre for Research on Bilingualism, and participants in CoNSALL 2018 for helpful discussion and comments. This work was supported in part by the US National Institutes of Health under NIDCD Grants T32DC000041 (GMO) and R01 DC010366 (EP, LB, ZG). Correspondence to Gary M. Oppenheim, School of Psychology, Bangor University, Bangor, Gwynedd LL57 2AS, UK. Email: g.m.oppenheim@bangor.ac.uk

Intro

How do a person's communicative abilities change with experience? The traditional generative view is that they do not. Most models of adult language production, for instance, aim to characterise the functioning of a mature expert system, where processing occurs without durably modifying its mechanisms (e.g. Dell, 1986; Levelt, Roelofs, & Meyer, 1999; see Roelofs, 2018, for an explicit rejection of persistent modification). Such models reflect a tradition wherein it has long been assumed that children acquire abstract knowledge of a language's syntax, morphology, and phonology by an early age (e.g. as a collection of rules, e.g. Pinker, 1991), and subsequent language development is limited to collecting new words. The underlying assumption is that acquisition quickly produces stable mastery of the target language. This assumption is perhaps most explicit in Chomsky's (1965 et passim) framework, which posits a distinction between a speaker's competence and their performance: even young children are assumed to have knowledge of what their language should be, knowing the rules of their language in the same way that a person might 'know' a particular word (competence), so difficulties lie in applying that knowledge to particular tasks (performance). When the same framework is applied to the case of learning multiple languages, a person's first language takes on a special status: whatever language they learn first should continue to be their strongest language throughout their life, meaning, *inter alia*, that it should always be easier for them to speak their first language than any that they acquired later.

Work in the past three decades, however, has brought increasing focus on language plasticity. Empirical work has inspired models in which experience-driven changes throughout the language system lead to changes in the relative availability of syntactic structures (Chang, Dell, & Bock, 2006), phonotactic combinations (Dell, Reed, Adams, & Meyer, 2000), and words (Oppenheim, Dell, & Schwartz, 2010). Contra the competence-performance approach noted above, the simple idea behind these new models is that the same kind of domain-general experience-driven incremental learning that supports a speaker's initial acquisition of a language also drives continued acquisition and optimisation throughout the lifespan (i.e. language development never stops; e.g. Seidenberg & MacDonald, 1999). In other words, the persistent priming effects that researchers report from laboratory experiments with adults reflect the same mechanism of change that supports early language acquisition. Applied to the question of bilingual production, this 'incremental learning' approach further predicts that, instead of simply prioritizing languages in terms of a discrete age of acquisition, the balance between languages should be sensitive to changes in how much a person uses them.

Indeed, extensive work on second language acquisition and education indicates that a person's ability in a second learned language (L2), if spoken by the majority of their society, often surpasses that in their first (L1; see Oller, Jarmulowicz, Pearson, & Cobo-Lewis, 2011, for a recent review). For instance, as a native Spanish-speaking child enters an environment where their L2, English, is used more often, their L2 can become their best, preferred, most fluent, or 'dominant' language in some if not all contexts (e.g. Dunn & Fox Tree, 2009; Flege, MacKay, & Piske, 2002; Jia & Aaronson, 2003). Although researchers have defined language dominance in many ways (for recent reviews, see Hamann, Rinke, & Genevska-Hanke, 2019, and Treffers-Daller, 2019), including bilinguals' estimates of their own competence (Kohnert, Bates, Hernandez, & Diego, 1999), measures of processing speed (i.e. performance) arguably offer better predictive validity (Flege et al., 2002). Moreover, consistent with the idea of speakers learning each time they use a language, dominance changes are manifest not only in children eventually knowing *more* words in their L2 than their L1, but also being able to access well-known words more easily (e.g. *dog* outpacing *perro*; Kohnert et al., 1999; Lambert, 1955; Mägiste, 1979). In fact, although estimates of *cumulative* language experience are clearly useful in predicting measures of language dominance (De Cat, 2019; Unsworth, Chondrogianni, & Skarabela, 2018), providing basic support for an incremental learning account, the additional contribution of *current* language experience (Bedore et al., 2012; Unsworth, 2013) suggests that dominance weights recent experiences more heavily than those in the distant past, and may therefore be better characterized as reflecting active adaptation rather than a more passive accumulation of experience.

Most reports of shifts in language dominance come from cross-sectional data (e.g. Kohnert et al., 1999; Mägiste, 1979 et passim), which is to say testing different children at different ages, assuming that each individual will follow the trajectory suggested by the combined data from the whole group. While it is often risky to draw individual conclusions from group data, logic requires only the highly plausible assumption that an L2-dominant sequential bilingual was L1-dominant early in L2 acquisition. Also, consistent with the idea that dominance shifts could occur within individuals, longitudinal studies of minority language speakers often suggest greater increases in L2 vocabulary than in L1 over time (e.g. Hoff & Ribot, 2017). Consistent with incremental learning accounts, language exposure estimates have been associated with the trajectories of several measures in each language, including performance-based estimates of auditory word recognition (Hurtado, Grüter, Marchman, & Fernald, 2014), but to our knowledge longitudinal associations with word production have been limited to competence-based estimates of vocabulary size (Hoff et al., 2012; Winsler, Diaz, Espinosa, & Rodriguez, 1999).

To close this 'performance' gap, we report a within-subjects longitudinal approach to assessing shifts in language dominance, using timed picture naming to track changes in how quickly and accurately 139 native and early-sequential Spanish-English bilingual children retrieve familiar words in each of their languages. As part of a larger study to evaluate changes in short-term learning in bilingual children with and without language impairment, we tested each child in each language, once per year, for up to four years, thus allowing us to assess, for the first time, within-subjects changes in language dominance via a performance-based measure of word production. We also tracked vocabulary growth via an untimed test (EOWPVT-3), thereby approximating previous competence-based approaches. Structured interviews of parents and teachers also provide estimates of each child's language exposure, at home and school, respectively, further allowing us to correlate changes in language accessibility and dominance with changes in language exposure. If changes in language dominance are driven by an individual's experience, then estimates of their experience should predict at least some of the variation in language dominance trajectories. Because these children all grew up in Spanish-dominant households and subsequently began attending monolingual English or bilingual primary schools, a mismatch between a child's home and school linguistic environments indexes a major longitudinal change in their linguistic experience. We therefore show not only that individual children change in their relative abilities in the two languages, but also assess whether their rate of change correlates with this discrepancy between their language environments.

Finally, it is important to note that much of focus on shifts in language dominance has been inspired by concern that learning or improving an L2 could reduce a child's ability to use their L1 (Köpke & Schmid, 2004), or that maintaining a first language could impair second language learning. Picture naming speed and accuracy are recognized as a particularly useful means of assessing such difficulties (Schmid & Köpke, 2009). Theoretically, competitive unlearning, on-line competition between language representations, or both could mean that any improvements in L2 access necessarily come at the cost of reducing access to one's L1 and attempts to maintain L1 could slow acquisition of L2. Although this study cannot address the theoretical question regarding mechanisms of language 'competition', it will allow us to address the question of absolute harm.

Method

Overview

We collected these data as part of a larger project aimed at characterizing and screening for possible language impairments among bilingual children; other aspects of this project will be reported separately. Methods for the core experimental procedures are described in this document; methods for supplementary assessments can be found in Appendix S1 or in separate articles where referenced as such.

Blocked cyclic picture naming is a widely used paradigm with adults (e.g. Damian, Vigliocco, & Levelt, 2001; Fink, Oppenheim, & Goldrick, 2018; Schnur, Schwartz, Brecher, & Hodgson, 2006) and has been used with children several times before (Charest, 2017; Ladányi & Lukács, 2016). Researchers use speed and accuracy to assess semantic context effects as participants simply try to name pictures that are repeatedly serially presented in related or unrelated blocks. When focusing on context effects, researchers often discard data from the first cycle of each block or treat it as a baseline where the context has not yet been established. At present, however, we simply use blocked cyclic naming as an example of a timed picture naming task, focusing specifically on the first cycle of each block to assess the baseline accessibility of picture names within each language.

Participants

For the larger study, over 300 children from the Austin Texas metropolitan area were recruited via their schools, from Kindergarten and Grades 2 and 4, and tested once per year for up to four years. Because the present report is concerned with shifts in language dominance over time, we include only the 139 Spanish/English bilingual children (74 female, 65 male; see Table 1)) for whom we had two or more years of data and whose parents initially reported more than 50% Spanish use at home (see Home Language Use questionnaire below). Thus, all grew up in Spanish-dominant homes and most also included some amount of English (mean estimated English proportion: 0.234; 75th/25th quartiles: .350/.106); no other languages were spoken at home. All reported normal or corrected-to-normal vision and hearing. Although we have not excluded data from children with suspected language impairments, all claimed results remain if doing so.

Table 1. Participant counts and descriptions by grade.

		Year of first test				L2 proportion	
Test grade	Total	K	2nd	4th	Age	Home	School
K	40	40[1]	0	0	5.81 (0.38)	0.26 (0.14)	0.29 (0.12)
1st	56	56[2]	0	0	6.85 (0.35)	0.26 (0.19)	0.41 (0.14)
2nd	89	41	48[1]	0	7.86 (0.31)	0.32 (0.18)	0.43 (0.15)
3rd	88	28	60	0	8.98 (0.34)	0.40 (0.21)	0.63 (0.20)
4th	56	0	36	20	9.90 (0.45)	0.42 (0.19)	0.66 (0.20)
5th	43	0	23	20	10.89 (0.44)	0.41 (0.17)	0.80 (0.21)
Total	139	59	60	20			

Table 1: Participant counts and descriptions by grade. [1] In the first year of the study, some participants completed a different form of the blocked cyclic naming task, so their first-year data are not directly comparable. [2] Three datasets were excluded due to equipment errors.

Language environment questionnaires

Home language use

This questionnaire was orally administered each year, via telephone, to each child's parent or guardian, in whichever language they preferred; the respondent sometimes differed from year to year. It included summary subjective estimates of specific language abilities (Inventory to Assess Language Knowledge; Peña, Gutierrez-Clellen, Iglesias, Goldstein, & Bedore, 2018) and age of first exposure, and, most important for the current analyses, hourly estimates of the child's input and output use of each language during out-of-school hours (included in Appendix S1, based on the Bilingual Input Output Survey; Peña et al., 2018).

School language use

This written questionnaire was completed each year by each child's schoolteacher. The respondent therefore typically differed from year to year. It included half-hourly estimates of the input and output use of each language (included in Appendix S1), typically assessed for the class as a whole.

Analyses

To estimate the discrepancy between each child's language environments—with which we later predict the trajectories of their changes in language dominance—we subtracted their home estimates of language use from their school estimates. That is, although these questionnaires potentially provide a much richer dataset, for our main analyses we simply use their first-year home questionnaire to estimate each child's mean proportion of English versus Spanish use during out-of-school hours (i.e. averaging the highly-correlated input and output estimates, after Bedore et al., 2012), and likewise use the first school questionnaire to estimate their mean proportion of English versus Spanish use during in-school hours; focusing specifically on either input and output does not change the significance of any reported result. These selections were made a priori, on theoretical bases, without considering other possible associations.

Blocked cyclic picture naming

Materials

Stimuli were 40 color images, depicting concrete nouns (e.g. *dog*, *shoe*), purchased from stock photo libraries. To minimize cross-language transfer, we used different stimuli in each language: 4 exemplars from each of 4 semantic categories, plus 4 unrelated objects. We selected these images for their high within-language name agreement among even the youngest children: our preliminary norms showed that Spanish-English bilingual kindergartners used their dominant names in 93.5% of cases for the English items, and 94.8% of cases for the Spanish items. The names of five items in each language can be considered cognates, but excluding them does not change the significance of any reported result.

Design

The naming task followed a standard blocked cyclic naming design. After an initial familiarization phase, where each picture appeared once in a random order, each picture appeared in one four-item block where every object was an exemplar of the same semantic category (e.g. *dog*, *snake*, *penguin*, *elephant*), and one four-item block where every object was an exemplar of a different semantic category (e.g. *dog*, *sock*, *apple*, *eggs*). Within each block, each set of items was presented six times in a pseudo-random order. The order of blocks was itself blocked by condition, and counterbalanced across participants, using 32 stimulus lists that further counterbalanced the order of semantic categories and items within each block. To maximize sensitivity to within-participant changes, each child received the same stimulus list each year.

To minimize cross-language transfer, order effects, and language switch costs, each language was tested in a separate session that was administered entirely in the target language. Thus, each session created a supportive environment for the target language, and no language mixing occurred within any part of an experiment session. These sessions also occurred on different days whenever possible (93.1% of cases), with their order counterbalanced across children (testing Spanish first in 49.4% of different-day sessions).

Within-session Procedure

Except for 51 (7%) sessions carried out within a lab setting, all children were tested in the quietest, most isolated location that their school could provide, using a laptop computer for stimulus presentation and response recording.

Familiarisation. At the start of each testing session, children saw each picture once and heard its auditorily presented name, which they then repeated aloud. They were instructed to use these names for the rest of the experiment.

Testing. During testing, children were instructed to name each picture as quickly and accurately as possible as soon as it appeared. Each trial began with a central fixation point displayed for 500ms, then a 250ms blank screen, and finally a single stimulus image in the center of the screen. The stimulus remained for 3500ms or until 1200 ms after the amplitude based voicekey detected a response. A red frame flashed around the stimulus picture when a response was detected to allow the experimenter to monitor voicekey sensitivity. The next trial began 750ms later. If no response triggered the voicekey, a message appeared instructing the experimenter to adjust the microphone or for the participant to speak up. The testing phase for each language consisted of 10x24-trial blocks, lasting approximately 20 minutes.

Apparatus

A 13" MacBookAir laptop computer controlled by MatLab 2010a with PsychToolbox extensions (Brainard, 1992; Pelli, 1997) presented all stimuli and digitally recorded responses via a SteelSeries Siberia V2 Full-Size Gaming Headset.

Response coding

Naming latencies were calculated online, and confirmed offline. Verbal responses were transcribed offline by speakers of the relevant languages. Following the standard approach in experimental psychology, we consider only the first response in each trial, disregarding any corrections that might follow. Responses were therefore coded as correct if children produced them without audible hesitation or correction, and they differed from the target response (as pronounced during familiarization) or a synonym (*bici* for *bicicleta*) by ≤ 1 phoneme. With this emphasis on the speed and accuracy of a child's first response, the assessment for this task thus focuses on performance, as is standard in experimental psychology.

Expressive One-Word Picture Vocabulary Test – Third Edition (EOWPVT-3)

The EOWPVT-3 for English (Brownell, 2000) and the Spanish-Bilingual Edition (Brownell, 2001) are norm-referenced tests of expressive vocabulary. The former is a 170-item picture naming test and the latter includes a subset of the same items. Unlike the blocked cyclic naming test, both are untimed, concerned only with a child's unfamiliarized accuracy. It is thus a test of vocabulary size rather than accessibility, and somewhat closer to the linguistic definition of 'competence'. Items are ordered based on difficulty in English.

Procedure

We administered the EOWPVT following the standard protocols, but continued testing for 14 items beyond the standard ceiling of six sequential incorrect responses.

Analyses

Accuracy was assessed via standard metrics for the EOWPVT, classifying responses as correct if a child produced an acceptable one-word description of the object in the target language, or a standard phonological variant thereof. Thus, *doggy* would be an acceptable variant of *dog*, but *cappertillar* would not be an acceptable variant of *caterpillar*. We consider raw naming scores, to the standard ceiling, as indicators of absolute vocabulary size. Note that, unlike the blocked cyclic picture naming procedure, accuracy assessments for the EOWPVT allow children to self-correct their responses (*e.g.* scoring "cappertillar… I mean caterpillar" as correct) and thus better approach the linguistic goal of assessing competence.

General analytical approach

All analyses use forms of mixed effects regression with maximal random effects for subjects and items, via the lme4 package in R. Error analyses apply logistic regression after excluding equipment errors and non-responses. Naming latency analyses use linear regression of inverse-transformed response times (-10000 *1/RT), for correctly detected correct responses only. P values are estimated via the Wald approximation method. Data and code for these analyses are available at https://osf.io/s2fht/.

Results

Characterizing the population

We used parent and teacher interviews to characterize the changes in these initially Spanish-dominant children's language environments over time (Table 1). According to the parent-based estimates, children's rate of home English use increased from a mean of .26 in Kindergarten to .41 in 5th grade. In the same timeframe, teachers reported an increase in the classroom school use of English, from a mean of .29 in Kindergarten to .80 in 5th grade, with a particular increase in 3rd grade implementing a practice of transitioning to English dominance in classrooms by that stage. Figure 1 illustrates the distributions of participants' language use at home and school in their first year of participation. Importantly, although these proportions are positively correlated (R=.20, p=.019)—consistent with an educational approach of instructing students in languages that they actually speak—they are far from perfectly correlated. Their mismatch provides the opportunity to predict changes in language accessibility from changes in language experience, via the plausible assumption that a child's maturation involves a transition from home-like linguistic environment to one where the school-like environment represents approximately half of their language exposure.

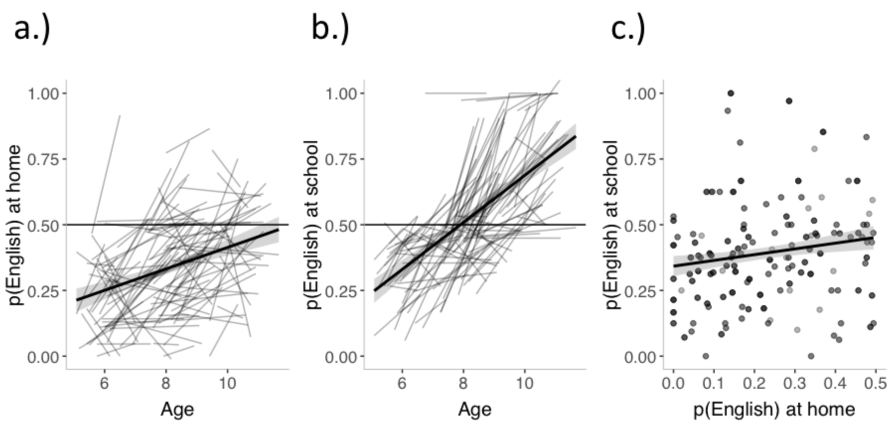

Figure 1: According to parent and teacher estimates, participants' use of English, relative to Spanish, increased over the course of the study, both at home (a) and at school (b). For individual participants, however, there was relatively little correspondence between the initial point estimates for the two settings (c). In Panels (a) and (b), each light grey line represents the estimated trajectory for one participant; the black line represents the estimated trajectory for the group as a whole.

Within-child changes in speed and accuracy in timed picture naming

For these 139 children, we have 366 blocked cyclic picture naming sessions in Spanish and 359 in English. As discussed earlier, we consider only the first naming cycle within each block, both to avoid the semantic effects that are usually the focus of the paradigm and because stronger repetition priming for weaker items (e.g. Griffin & Bock, 1998; Oppenheim, 2018) would attenuate differences that stem from baseline accessibility. This provides 27,588 observations for the mixed effects logistic regression of error data, and 24,889 correct responses for the mixed effects linear regression of naming latencies, illustrated in Figure 2a and b respectively.

To assess experience-driven changes in language accessibility we fit the same mixed effects regression model to both the error rates and the naming latencies, including by-subjects and by-items maximal random effects and the following three fixed effects and their interactions:

1. **Age**, denoting the child's age in years at the time of testing. To ease interpretation, this continuous variable is centered around 6 years old, the mean age of our youngest cohort.
2. **Language,** a binary-coded contrast where the child's L1, Spanish, is coded as 0 and their L2, English, is coded as 1.
3. **Home-to-school difference in L2 use,** a continuous variable that is centered around its approximate midpoint, .15. As illustrated in Figure 1c, this is the initial difference between each child's teacher-estimated rate of English use in their classroom and their parent-estimated rate of English use at home. We use the Year 1 estimates for these values (rather than variable per-year estimates) for two reasons: 1.) they best represent the initial discrepancy between the child's language environments, and 2.) they allow us to statistically test the association between changes in language use and changes in language production ability as a three-way interaction between this predictor and Age and Language.

With this coding, the intercept for a fitted model provides its subject- and item-variability-adjusted estimate for a 6-year-old child naming in Spanish, and main effects and interactions represent deviations from that baseline. Results of these analyses are given in Table 2.

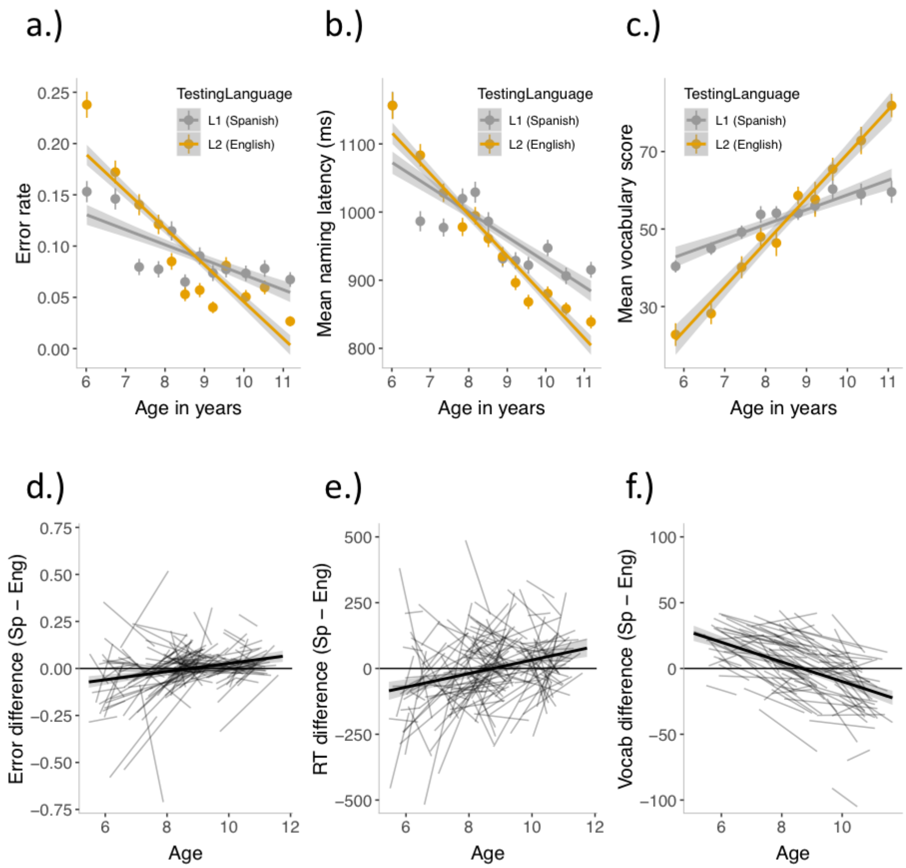

Figure 2: Changes in language dominance as a function of age, as assessed by (a) error rates and (b) naming latencies in the blocked cycling naming task, and (c) a vocabulary score from the EOWPVT. In the second row (d,e,f), each light grey line represents the cross-language differential trajectory for one participant; the black line represents that for the group as a whole.

	Errors 27588 obs: 139 subjects, 40 items					Naming latencies 24889 obs: 139 subjects, 40 items					Vocabulary test 707 obs: 138 subjects			
	β	SE	t	p	OR	β	SE	t	p	ms	β	SE	t	p
Intercept	-2.087	0.170	-12.25	$<.001$	0.120	-10.301	0.291	-35.42	$<.001$	970.8	41.07	1.06	39.17	$<.001$
Language	0.657	0.239	2.74	.006	1.929	0.918	0.287	3.2	.001	95.0	-19.89	1.92	-10.38	$<.001$
Age	-0.188	0.049	-3.84	$<.001$	0.828	-0.661	0.095	-6.96	$<.001$	-58.5	4.31	0.35	12.39	$<.001$
Home-school diff	0.975	0.447	2.18	.029	2.650	-1.405	1.058	-1.33	.18	-116.5	1.84	4.60	0.40	.69
Lang X Age	-0.344	0.065	-5.26	$<.001$	0.709	-0.339	0.065	-5.21	$<.001$	-32.5	7.47	0.53	14.20	$<.001$
Lang X Diff	0.786	0.613	1.28	.20	2.194	1.324	0.724	1.83	.068	129.7	-7.04	8.25	-0.85	.39
Age X Diff	0.005	0.179	0.03	.98	1.005	0.662	0.408	1.62	.11	-116.5	-3.22	1.50	-2.15	.032
Lang X Age X Diff	-0.415	0.231	-1.79	.073	0.661	-0.563	0.252	-2.23	.026	-55.4	4.69	2.22	2.11	.035

Table 2: Regression results for the blocked cyclic naming experiment and EOWPVT. Intercept: Language = L1 (Spanish), Age = 6 years old, Home-to-school difference in L2 use = 0.15. Language: L2 (English). Age (in years over 6). Home-to-school difference in L2 use (centered around 0.15)

The youngest children were approximately ten percent slower (Language: β=0.918, p=.001) and twice as likely to err (β=0.657, p=.006) when naming pictures in English compared to Spanish, thus confirming their initial questionnaire-based characterization as Spanish-dominant. Their Spanish naming also significantly improved with age, growing faster (Age: β=-0.661, p<.001) and more accurate (β=-0.188, p<.001). Their English, however, improved more quickly, both in terms of speed (Language X Age: β=-0.339, p<.001) and accuracy (β=-0.344, p<.001). These model parameters thus estimate an average shift from L1 dominance to L2 dominance around 6+0.918/0.339 = 8.71 years if assessed in terms of speed, or around 6+0.657/0.344 = 7.91 years if assessed in terms of accuracy. This not only replicates the crossover in language dominance reported in previous cross-sectional studies, but further demonstrates that the shift actually occurs within individuals.

A within-subjects longitudinal approach also allows us to examine how individual differences in experience affect language learning. If our observed changes in language dominance are actually driven by changes in a child's language environment, then the Language X Age interaction should be stronger for those children whose home and school environments differ more. And we see evidence of this interaction in both naming latencies (Language X Age X Home-to-school difference in L2 use: β=-0.563, p=.026) and error rates (albeit less consistently: β=-0.415, p=.073). Thus, the observed changes in language dominance can be linked to changes in language experience.

Converging evidence from vocabulary tests

Other tasks and measures can provide converging evidence through more traditional means. Instead of accessibility, the Expressive One-Word Picture Vocabulary Test aims to assess lexical competence by using unfamiliarized, untimed, picture naming to estimate the size of a child's vocabulary in each language. Applying the same mixed effects linear regression models to these scores yields similar within-child results (also listed in Table 2).

The youngest children's vocabulary estimates were significantly greater in Spanish than English (Language: β=-19.89, p<.001). Their Spanish naming scores also significantly improved with age (Age: β=4.31, p<.001). Their English naming scores, however, improved more quickly (Language X Age: β=7.47, p<.001), thus estimating a change from L1 dominance to L2 dominance around 6+19.89/7.47 = 8.66 years of age, converging with the estimates from the timed naming task. Finally, this interaction was again stronger for those children whose home and school environments differed more (Language X Age X Home-to-school difference in L2 use: β=4.69, p=.035), thus once again linking the observed changes in language dominance with changes in language experience.

Discussion

According to contemporary understanding of how language development generally works (e.g. Chang et al., 2006; Dell et al., 2000; Oppenheim et al., 2010; Seidenberg & Macdonald, 1999), there is little question that language abilities *should* correlate with practice. Rather than merely prioritizing one's languages by age of acquisition, incremental learning accounts (Oppenheim et al., 2010) predict that a speaker should benefit from each language experience: retrieving 'dog' should strengthen links for retrieving 'dog' in the future, retrieving '*perro*' should strengthen links for retrieving '*perro*', and as a speaker transitions from an L1-dominated environment to an L2-dominated environment these basic practice effects should accumulate into changes in their language dominance. The question is whether they actually do. In this paper we have empirically demonstrated within-speaker shifts from L1-dominance to L2-dominance, as assessed via both performance and competence measures of word production. Our consideration of speed and error rates in a timed picture naming task represents a psychological emphasis on performance and fluency, while consideration of attainment in an untimed vocabulary test better represents a linguistic emphasis on competence, and both approaches provide converging evidence for these dominance shifts. We have further shown that these changes correspond to gradient changes in the speakers' linguistic environments, reinforcing the causal link between ongoing experience and demonstrated language dominance.

This general result is consistent with an extensive literature demonstrating links between estimates of bilingual children's current (Bedore et al., 2012) and cumulative (De Cat, 2019) language experience and use and measures of both their absolute and relative language abilities (see Unsworth, Chondrogianni, & Skarabela, 2018, for a recent review). But to our knowledge, this is the first longitudinal evidence of long-term shifts in language dominance, as assessed via performance measures of word production. Given the abundance of previous studies of bilingual language acquisition, it might seem odd that such evidence has not previously been reported. Reasons may include the well-known difficulty in conducting longitudinal research (Oller et al., 2011), the common expedient of testing only one language, emphasis on unspeeded tests to measure proficiency or identify disorders, and the use of language-switching and translation tasks to maximize and study cross-language interference (e.g. Kohnert & Bates, 2002; Kohnert et al., 1999).

One potential concern is that, because we tested children in their schools, where English tended to be more dominant than in their homes, our data may reflect temporary context-dependent changes in accessibility rather the grander context-independent changes that we have claimed. Although we agree that memory retrieval should generally incorporate context as a cue, we can offer several points to support our stronger interpretation. First, although schools varied in their amount of Spanish use, all used some and could therefore be considered bilingual contexts, thus providing environmental support for both languages (e.g. Grosjean, 2001). Second, each session was run entirely in its target language, thereby creating a maximally supportive context for it (e.g. Wu & Thierry, 2010). Third, according to most thinking in the wider memory literature (see e.g. Smith & Vela, 2001, for a review), context is typically thought to support weak associations rather than disrupt access to overlearned associations, such as words that a speaker has used hundreds or thousands of times. Finally, for the blocked-cyclic naming task, pre-experiment familiarisation provided an opportunity to re-establish the pictures' names in the target language, which is conservative with respect to our claims of a shift in language dominance. That is, although some have suggested that speakers can easily overcome the effects of language attrition (e.g. Köpke & Genevska-Hanke, 2018)—via processes that experimental psychologists might identify with repetition priming (Scarborough, Cortese, & Scarborough, 1977) and error-proportional incremental learning (Oppenheim et al., 2010)—any resulting attenuation of L1 attrition would both make it more difficult detecting a change in dominance and be insufficient to explain children's continued L1 improvement in both tasks and all measures.

Finally, we note that although we have claimed shifts in language dominance as a whole, we have considered only a small component of language: meaning-driven word production. Language dominance is a multidimensional construct (e.g. Hamann et al., 2019; Treffers-Daller, 2019), and there is mounting evidence that the speed of word retrieval may be a particularly sensitive or labile measure of it. In contrast, measures of receptive language speed and accuracy often show smaller or later changes in dominance, or no change at all (see Oller et al., 2011, for review). Such observations are consistent with continuous learning, context-dependency, and non-monolithic views of language processing.

Attrition versus co-improvement.

One remarkable feature of our results is that the change from L1 to L2 dominance cooccurred with absolute improvements in L1 abilities. That is, speakers' speed, accuracy, and vocabulary estimates increased in their L1 during the period of study; they simply increased more dramatically in their L2. The considerable research focus on L1 attrition and suppression (e.g. Köpke & Schmid, 2004; Linck, Kroll, & Sunderman, 2009; Steinhauer, this volume) can lend the impression that language learning is a zero-sum game, where improvement in a L2 necessarily harms the L1, so our finding of co-improvement might seem like a surprising and even contradictory result. (As noted above, although our pre-experiment familiarization in the blocked cyclic naming task could have attenuated the effects of L1 attrition, it would not explain the significant continued improvement in the L1 version of that task, nor in the L1 vocabulary assessments, which proceeded without familiarization.) But in fact, both attrition and co-improvement are neatly explained by computational models of incremental language learning that include some mechanism for forgetting or 'unlearning'.

For instance, in a recent computational project, Oppenheim, Dell, & Schwartz (2010) applied a classic incremental learning algorithm (Rescorla & Wagner's, 1972, 'delta rule', originally proposed to describe operant conditioning in rats) to the task of mapping from meaning to words in language production. For each retrieval, they activated several of the network's semantic features ('mammalian', 'terrestrial'), and it attempted to map them onto an appropriate word (*dog*). After each attempt, its learning algorithm simply strengthened the links from the activated semantic features to the intended word (*dog*), and weakened the links from those features to any erroneously activated words (*bat*, *whale*), implementing stronger weight changes when the activation patterns had been further from correct (i.e. error-proportional learning). Although the network was not 'born' knowing any of the correct mappings, through this incremental learning algorithm it grew able to correctly activate *dog* when it encountered 'mammal' and 'terrestrial', *whale* when it encountered 'mammal' and 'aquatic', and *car* when it encountered 'vehicular' and 'terrestrial'.

More relevant to the current point, when this model was run on a blocked cyclic naming experiment—the same task that we used here, where participants repeatedly name a small set of semantically related or unrelated pictures—its incremental learning algorithm created both 1.) facilitation when the model tried to name the same picture later (*dog*... dog), thus mapping from the same semantic features to the same word, and 2.) interference when it tried to retrieve a semantically related name instead (*whale*... dog), thus mapping some of the same semantic features to a different word. When the model performed a naming task without repetition (*bat, whale,* dog), the unused words in its lexicon grew less accessible each time it retrieved a word with a similar meaning, so it seemed to be losing access to those words, thus resembling L1 attrition. But when the task introduced repetition (*dog, bat, whale,* dog), the resulting facilitation outweighed the interference, thus generating net improvement and illustrating how the same learning process can generate both increases and decreases in performance, depending on the ratios of relevant experience.

Although this model was originally proposed to account for within-language effects, like cumulative semantic interference, extending it to multilingual production requires only a few assumptions (such as mechanisms for language-based lexical activation and/or selection) to explain why retrieving *dog* interferes with *gato* as well as *cat* (see Runnqvist, Strijkers, Alario, & Costa, 2012, for relevant discussion). In the same way that it explains both semantic interference and repetition priming effects in blocked cyclic naming, the model can explain both cases of both L1 attrition and co-improvement as differences in the ratio of *cats* to *gatos*: trying to retrieve *cat* may cause weakening of some connections supporting *gato* (e.g. from the shared 'mammal' and 'terrestrial' semantic features), but later attempts to retrieve *gato* can initiate their repair and even improvement. This repair and improvement can be relatively rapid because the learning algorithm is error-proportional, meaning that it applies stronger corrections when its activation patterns are less consistent with its desired behavior. One underappreciated consequence of this error-proportionality is that, although the networks connections reflect the cumulative effects of all its previous experience—and can thus account for basic frequency effects as well as 'weaker links'-related phenomena (Gollan, Montoya, Cera, & Sandoval, 2008)—they disproportionately reflect its most recent experience. It can therefore account for both relatively rapid shifts in language dominance and attenuation of L1 attrition upon re-immersion in an L1 environment (Köpke & Genevska-Hanke, 2018), as well as the particular importance of bilinguals' current language use in predicting their relative abilities (Bedore et al., 2012; Unsworth, 2013). Thus, switching entirely to an L2 may lead to the gradual loss of an L1, but occasionally interleaving episodes of L1 use may be sufficient to not only undo the damage but even reverse it (cf. McCloskey & Cohen, 1989). This pattern may explain why reports of L1 attrition have typically focused on cases of L2 immersion, with minimal ongoing L1 use. Thus, although it is not developed to address the full range of sociological and motivational factors that may contribute to language dominance in humans, a strength of this computational modelling approach is that it illustrates how such domain-specific tendencies and variation can emerge from domain general principles.

Conclusion

Experience-driven plasticity is a core feature of all levels of language use and representation. Not only do we continually modify our representations and procedures within a language, dynamically learning and unlearning even well-established words, but we also modify our access to multiple languages.

This is because the same kinds of domain-general incremental learning processes operate at every level of the system, yielding experience driven changes. More generally, such effects reflect the kind of basic incremental learning processes that underlie even complex behaviors like language production.

References

Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., . . . Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism*, *15*(3), 616–629. https://doi.org/10.1017/S1366728912000090

Brainard, D. H. (1992). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Brownell, R. (2000). Expressive one-word picture vocabulary test. Novato, CA: Academic Therapy Publications.

Brownell, R. (2001). Expressive One-Word Picture Vocabulary Test—Spanish-Bilingual Edition. Novato, CA: Academic Therapy Publications.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272. https://doi.org/10.1037/0033-295X.113.2.234

Charest, M. (2017). Cumulative semantic interference in young children ' s picture naming. *Applied Psycholinguistics*, 1–19. https://doi.org/10.1017/S0142716416000461

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press. https://doi.org/10.21236/AD0616323

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), 77–86. https://doi.org/10.1037/e501882009-313

De Cat, C. (2019). Predicting Language Proficiency in Bilingual Children. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263119000597

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1355–1367. https://doi.org/10.1037/0278-7393.26.6.1355

Dunn, A. L., & Fox Tree, J. E. (2009). A quick, gradient Bilingual Dominance Scale*. *Bilingualism: Language and Cognition*, *12*(03), 273. https://doi.org/10.1017/S1366728909990113

Fink, A., Oppenheim, G. M., & Goldrick, M. A. (2018). Interactions between lexical access and articulation. *Language, Cognition and Neuroscience*, *33*(1), 12–24. https://doi.org/10.1080/23273798.2017.1348529

Flege, J. E., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, *23*(4), 567–598. https://doi.org/10.1017/S0142716402004046

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*(3), 313–338. https://doi.org/10.1006/jmla.1997.2547

Grosjean, F. (2001). The Bilingual's Language Modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing.* (pp. 1–22). Blackwell Publishing.

Hamann, C., Rinke, E., & Genevska-Hanke, D. (2019). Editorial: Bilingual language development: The role of dominance. *Frontiers in Psychology*, *10*(MAY), 1–3. https://doi.org/10.3389/fpsyg.2019.01064

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Hoff, E., & Ribot, K. M. (2017). Language Growth in English Monolingual and Spanish-English Bilingual Children from 2.5 to 5 Years. *Journal of Pediatrics*, *190*, 241-245.e1. https://doi.org/10.1016/j.jpeds.2017.06.071

Hurtado, N., Grüter, T., Marchman, V. A., & Fernald, A. (2014). Relative language exposure, processing efficiency and vocabulary in Spanish-English bilingual toddlers. *Bilingualism*, *17*(1), 189–202. https://doi.org/10.1017/S136672891300014X

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, *24*(01), 131–161. https://doi.org/10.1017/S0142716403000079

Kohnert, K. J., & Bates, E. (2002). Balancing Bilinguals II: Lexical Comprehension and Cognitive Processing in Children Learning Spanish and English. *Journal of Speech Language and Hearing Research*, *45*, 347–359. https://doi.org/10.1044/1092-4388(2002/027)

Kohnert, K. J., Bates, E., Hernandez, A. E., & Diego, S. (1999). Balancing bilinguals: lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research*, *42*, 1400–1413. https://doi.org/10.1044/jslhr.4206.1400

Köpke, B., & Genevska-Hanke, D. (2018). First language attrition and dominance: Same same or different? *Frontiers in Psychology*, *9*(NOV), 1–16. https://doi.org/10.3389/fpsyg.2018.01963

Köpke, B., & Schmid, M. S. (2004). Language attrition. In M. S. Schmid, B. Köpke, M. Keijzer, & L. Weilemar (Eds.), *First language attrition: Interdisciplinary perspectives on methodological issues* (pp. 1–43). Amsterdam: Benjamins.

Ladányi, E., & Lukács, Á. (2016). Lexical conflict resolution in children with specific language impairment. *Journal of Communication Disorders*, *61*, 119–130. https://doi.org/10.1016/j.jcomdis.2016.04.004

Lambert, W. E. (1955). Measurement of the linguistic dominance of bilinguals. *Journal of Abnormal and Social Psychology*, *50*(2), 197–200. https://doi.org/10.1037/h0042120

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(01), 1–38. https://doi.org/10.1017/S0140525X99001776

Mägiste, E. (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 79–89. https://doi.org/10.1016/S0022-5371(79)90584-X

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. *Psychology of Learning and Motivation*, *24*, 109–165. https://doi.org/10.1016/S0079-7421(08)60536-8

Oller, D. K., Jarmulowicz, L., Pearson, B. Z., & Cobo-Lewis, A. B. (2011). Rapid spoken language shift in early second-language learning. In A. Y. Durgunoglu & C. Goldenberg (Eds.), *Language and literacy development in bilingual settings* (pp. 94–120).

Oppenheim, G. M. (2018). The paca that roared : Immediate cumulative semantic interference among newly acquired words. *Cognition*, *177*(August 2018), 21–29. https://doi.org/10.1016/j.cognition.2018.02.014

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: a model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252. https://doi.org/10.1016/j.cognition.2009.09.007

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Peña, E. D., Gutierrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2018). Bilingual English Spanish Assessment (BESA). Baltimore: Brookes.

Pinker, S. (1991). Rules of Language. *Science*, *253*, 530–535. https://doi.org/10.1016/j.cub.2004.10.009

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variation in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: II. Theory and research* (pp. 64–99). New York: Appleton-Century-Crofts.

Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, *172*, 59–72. https://doi.org/10.1016/j.cognition.2017.12.007

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. https://doi.org/10.1016/j.jml.2012.02.007

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and Repetition Effects in Lexical Memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 1–17. https://doi.org/10.1037/0096-1523.3.1.1

Schmid, Monika S, & Köpke, B. (2009). L1 attrition and the mental lexicon. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 209–238). Multilingual Matters Clevedon.

Schnur, T. T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language*, *54*(2), 199–227. https://doi.org/10.1016/j.jml.2005.10.002

Seidenberg, M. S., & Macdonald, M. C. (1999). A Probabilistic Constraints Approach to Language Acquisition and Processing. *Cognitive Science*, *23*(4), 569–588. https://doi.org/10.1207/s15516709cog2304

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*(2), 203–220. https://doi.org/10.3758/BF03196157

Treffers-Daller, J. (2019). What Defines Language Dominance in Bilinguals? *Annual Review of Linguistics*, *5*(1), 375–393. https://doi.org/10.1146/annurev-linguistics-011817-045554

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism*, *16*(1), 86–110. https://doi.org/10.1017/S1366728912000284

Unsworth, S., Chondrogianni, V., & Skarabela, B. (2018). Experiential measures can be used as a proxy for language dominance in bilingual language acquisition research. *Frontiers in Psychology*, *9*(OCT), 1–15. https://doi.org/10.3389/fpsyg.2018.01809

Winsler, A., Diaz, R. M., Espinosa, L., & Rodriguez, J. L. (1999). When learning a second language does not mean losing the first: bilingual language development in low-income, Spanish-speaking children attending bilingual preschool. *Child Development*, *70*(2), 349–362. https://doi.org/10.1111/1467-8624.t01-1-00026

Wu, Y. J., & Thierry, G. (2010). Investigating bilingual processing: The neglected role of language processing contexts. *Frontiers in Psychology*, *1*(NOV), 1–6. https://doi.org/10.3389/fpsyg.2010.00178

Supporting Information

Appendix S1. Language Use Questionnaires