Bilingualism: Language and Cognition 15 (1), 2012, 102–116 © Cambridge University Press 2011 doi:10.1017/S1366728910000490

The receptive–expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms*

TODD A. GIBSON
School of Communication Sciences and Disorders, The University of Memphis
D. KIMBROUGH OLLER
School of Communication Sciences and Disorders, The University of Memphis
LINDA JARMULOWICZ
School of Communication Sciences and Disorders, The University of Memphis
CORINNA A. ETHINGTON
Department of Counseling, Educational Psychology, and Research, The University of Memphis

(Received: November 11, 2009; final revision received: June 29, 2010; accepted: October 10, 2010; First published online 7 September 2011)

Adults and children learning a second language show difficulty accessing expressive vocabulary that appears accessible receptively IN THEIR FIRST LANGUAGE *(L1). We call this discrepancy the receptive–expressive gap. Kindergarten Spanish (L1) – English (L2) sequential bilinguals were given standardized tests of receptive and expressive vocabulary in both Spanish and English. We found a small receptive–expressive gap in English but a large receptive–expressive gap in Spanish. We categorized children as having had high or low levels of English exposure based on demographic variables and found that the receptive–expressive gap persisted across both levels of English exposure. Regression analyses revealed that variables predicting both receptive and expressive vocabulary scores failed to predict the receptive–expressive gap. The results suggest that the onset of the receptive–expressive gap in L1 may have been abrupt. We discuss possible mechanisms underlying the phenomenon.*

Keywords: bilingualism, second-language learning, language attrition, vocabulary learning, school-age children

Research to be reviewed here shows that learners immersed in a second language (L2) score higher on tests of receptive vocabulary than expressive vocabulary IN THEIR FIRST LANGUAGE (L1). Even after controlling for the difference in difficulty between receptive and expressive vocabulary testing tasks, children appeared to have problems accessing their L1 expressive vocabulary. We term this systematically uneven relationship between receptive and expressive vocabulary the RECEPTIVE–EXPRESSIVE GAP.

The receptive–expressive gap in bilinguals' L1 is thus counterintuitive and warrants further investigation to assess factors that may influence its occurrence. A first step is a characterization of the phenomenon's robustness, the strong tendency for the receptive–expressive gap to occur in most children even if they have been immersed in L2 for only a relatively short period. The current work also considers speculations about causes of the receptive–expressive gap. Few researchers have focused on the receptive–expressive gap explicitly, but embedded in a variety of studies is evidence that this phenomenon is present across languages. Within this literature, the evidence of a receptive–expressive gap pertains to children who start school or preschool with L1 skills only, and begin L2 learning in kindergarten (K) or preschool.

* This research was supported by a grant from the National Institutes of Health, National Institue of Child Health & Human Development (R01 HD046947 to D. Kimbrough Oller, Principal Investigator), and by the Plough Foundation to D. Kimbrough Oller. Thanks to the anonymous reviewers who gave us useful feedback on an earlier version.

Address for correspondence:
Todd A. Gibson, 1 University Station A 1100, Austin, TX 78712, USA
todd.gibson@austin.utexas.edu

Background on the receptive–expressive gap

Muñoz & Marquardt (2003) did not directly investigate the receptive–expressive gap, but embedded within their study of normal bilingual language patterns was evidence of its existence. They tested Spanish–English bilingual adults whose L1 was Spanish but who had acquired both languages by age 10. Participants in the study took non-standardized vocabulary tests that implemented both picture pointing tasks (as a measure of receptive vocabulary) and picture naming tasks (as a measure of expressive vocabulary). The same set of 100 pictures was used for both tasks.

Results showed that the average number of pictures correctly identified and named in English was 99.15 and 90.85, respectively, a receptive–expressive gap of 8.3 correct answers in English. The receptive–expressive gap in Spanish was much higher, at 14.9 correct answers (95.8 for receptive compared to 80.9 expressive). If it were the case that the gap between receptive and expressive vocabulary only reflected the comparative ease of recognition tasks relative to recall tasks, one should find a similar ratio of expressive vocabulary to receptive vocabulary in BOTH languages, since the same picture stimuli were used across languages and tasks. However, the authors stated that "For bilingual speakers . . . Spanish may be associated with a more pronounced receptive–expressive discrepancy than is evident in English. Bilingual speakers may have difficulty retrieving words in Spanish that they understand" (Muñoz & Marquardt, 2003, p. 1124). They further reported that participants often made remarks suggesting lexical access difficulty, such as "I know what that is but I can't think of it" (p. 1124).

The receptive–expressive gap has been present not only in Spanish–English bilingual adults, but also in Spanish–English bilingual preschoolers (Miccio, Tabors, Paez, Hammer & Wagstaff, 2005), K children (Oller, Jarmulowicz, Gibson & Hoff, 2007), and older school-age children (Oller & Eilers, 2002; Oller, Pearson & Cobo-Lewis, 2007; Windsor & Kohnert, 2004). Preschool children of Puerto Rican descent and living in Pennsylvania or Massachusetts were given standardized tests of receptive and expressive vocabulary in both English and Spanish (Miccio et al., 2005). The tests had a mean of 100 and a standard deviation of 15. Results yielded a receptive–expressive gap of 5 standard points in English (L2) but 19 standard points in Spanish (L1).

Using the same set of standardized vocabulary tests as Miccio et al. (2005), Oller, Jarmulowicz, Gibson & Hoff (2007) examined vocabulary of Hispanic K children in Memphis, Tennesee. These English language learners had an average receptive–expressive gap of 6 standard points in English but 23 standard points in Spanish. The authors speculated that the receptive–expressive gap could have resulted from a suppression mechanism that might aid in the shift from L1 to L2 by limiting L1 interference during the acquisition of L2 vocabulary. Windsor & Kohnert (2004) used a similar set of standardized tests on Spanish–English sequential bilinguals age 8–13 years. The receptive–expressive gap in English was 5 standard points but 12 standard points in Spanish. Their smaller Spanish receptive–expressive gap may have been due to more strict participant criteria.

In a reanalysis of a study of both K and school-age Spanish–English bilinguals in Miami, Florida, Oller, Pearson and Cobo-Lewis (2007) found an English receptive–expressive gap in K, second grade and fifth grade of 0.4, –0.02, and 4 standard points, respectively. The Spanish receptive–expressive gap in K, second grade and fifth grade was 28, 28, and 19 standard points, respectively. These authors speculated that the receptive–expressive gap was associated with the children's decreased use of Spanish, resulting in difficulty producing Spanish vocabulary, while retaining relative ease of understanding Spanish vocabulary. An even larger gap was found by Swanson, Rosston, Gerber and Solari (2008). Again, using similar standardized vocabulary tests, Spanish–English bilingual third grade students who spoke Spanish at home had an English receptive–expressive gap of 6 standard points and a receptive–expressive gap of 53 standard points in Spanish.

The receptive–expressive gap is not restricted to Spanish–English bilinguals. Embedded within a study by Kan & Kohnert (2005) was evidence that Hmong-speaking children attending an English-only preschool presented with a receptive–expressive gap in Hmong (L1) that was larger than the receptive–expressive gap in English. As part of non-standardized vocabulary testing, children took picture pointing and picture naming vocabulary tests in Hmong and in English. Results showed that the raw score difference between receptive and expressive vocabulary in Hmong (28.4 items correct and 16.5 items correct, respectively) was statistically significant, while the difference between receptive and expressive vocabulary in English (24.5 items correct and 22.9 items correct, respectively) was not statistically significant. The authors reported that their results were "consistent with the notion that there is a general weakening in L1 lexical-semantic skills" Kan & Kohnert, 2005, p. 379) as children make the shift from Hmong to English dominance.

Yan & Nicoladis (2009) investigated the discrepancy between L1 receptive and expressive vocabulary in French–English bilingual school-age children in an English-dominant area of Canada. During a non-standardized picture naming task, bilingual children not only named pictures, but also reported whether they knew the target word when they were unable to name the picture (being unable to name a known word was considered a "tip of the tongue" state). Bilingual children then were given a follow-up comprehension test to determine whether they recognized the tip-of-the-tongue words. The same assessment was performed with a monolingual comparison group.

Results of the Yan & Nicoladis (2009) study revealed that the bilingual children named significantly fewer pictures than the monolingual children, and the bilingual children more frequently reported tip-of-the-tongue states. Of the pictures the children were unable to name immediately, about 70% were correctly identified during the follow-up comprehension task. The authors speculated that the naming difficulty experienced by bilingual children might have been due to interference from the second language, and the cultural context in

which the bilinguals lived (English-dominant area of Canada) may have played a role as well.

The study of bilingual vocabulary knowledge has thus repeatedly provided evidence of a receptive–expressive gap in L1, but in most cases the evidence has been embedded in studies that focused on different issues. Little attention has been paid to the existence of the gap and even less to possible factors that may influence it. Of the articles cited, only Oller, Jarmulowicz, Gibson and Hoff (2007) speculated on a possible mechanism for the gap. The present research begins with an exploration of possible influencing factors and concludes with speculations on mechanisms.

Possible influences on the receptive–expressive gap

Many factors influence vocabulary development in both monolingual and bilingual children. These include (but are not limited to) the quantity, quality, and source of language input (Hoff, 2006; Kan & Kohnert, 2005). Other factors being equal, amount of input in each language plays a role in children's vocabulary size (Hoff, 2006; Pearson, Fernández, Lewedeg & Oller, 1997). Children in vocabulary-rich environments are more likely to develop extensive vocabularies than are children in vocabulary-impoverished environments (Hart & Risley, 1995). Furthermore, the number of people in the home and their relationship to the child influences both the quantity and the quality of language input (Hoff, 2006; Morales & Hanson, 2005).

The factors that influence vocabulary development in bilingual speakers may or may not also influence the appearance or degree of the receptive–expressive gap. An evaluation of possible factors is important in order to gain perspective on the robustness of the receptive–expressive gap and to provide a perspective from which speculations about underlying mechanisms can be constructed. Our research provided the opportunity to explore nine predictions regarding influences on the gap. Each of the predictions was formulated based on demographic information available to us in an ongoing investigation of young second-language learners.

The present study: Predictions and rationale

We reasoned that increases in English exposure would have at least two effects. First, they would directly impact vocabulary scores in both English and Spanish, with English scores increasing in the course of English exposure and Spanish scores decreasing. Second, we reasoned that as English exposure increases, the magnitude of the receptive–expressive gap in Spanish would increase. The nine predictions related to factors that could influence the receptive–expressive gap were all based on expectations derived from questionnaire data suggesting that English exposure would likely have been higher in one subgroup of the children than another. We reasoned as follows:

1. A child born in the USA likely would have had more exposure to English (and would thus have had more time to develop a receptive–expressive gap in Spanish) than a child born in a Spanish-speaking country.
2. Children who had attended preschool where English had been spoken at least half the time should have had more English exposure and thus larger receptive–expressive gaps in Spanish than children who had not attended preschool in English.
3. We defined the age of first exposure to English as the time when either the child began preschool in English or when the child was first regularly exposed to English on a frequent basis as reported in the questionnaire. Children exposed to English before age four were likely to have experienced more English input than children exposed to English at a later date. Therefore, we predicted that children exposed to English before age four might have a larger receptive–expressive gap than children exposed to English at a later age.
4. Mothers of children who had resided in the USA for more than three years would have had more time to learn English than mothers who had been here for a shorter time. Thus we predicted that children of mothers with longer residency in the USA might have been exposed (through their mothers) to more English than the children of short-residency mothers, and might thus show larger receptive–expressive gaps.
5. Mothers who reported that they had some English language proficiency may have been more likely to speak English to their children than mothers who reported they did not have any English language proficiency. Therefore, we anticipated that the children of mothers with some English language proficiency would have larger receptive–expressive gaps than the children of mothers without English language proficiency.
6. Hakuta & D'Andrea (1992) offered empirical evidence to justify using Hispanic mothers' level of education as a proxy for English language proficiency. We compared mothers with relatively lower levels of education (six years or less) with mothers with relatively higher levels (seven years or more). We predicted that the children of mothers with higher levels of education might present with the larger receptive–expressive gap, since they might have been exposed to more English than the children of mothers with fewer years of education.

7. When more than two adults lived in the homes of the children in our study, these adults were almost always extended family members who spoke Spanish primarily or exclusively. We predicted that increases in the number of adults in the home would increase the amount of Spanish spoken in the home. Therefore, we expected children in homes with only one or two adults to be exposed to more English (and therefore they might present with larger receptive–expressive gaps) than children in homes with three or more adults.
8. On the other hand, we assumed that when children were in the presence of other CHILDREN in the home, the likelihood would be higher for them to speak to each other in English. Therefore, children living in homes with more than two children (including the research participant) would be expected to be exposed to more English and might show a greater receptive–expressive gap than children who lived in homes with only one or two children.
9. Shin (2002) found a birth-order effect: First-born bilingual children spoke their parents' L1 more often than second-born children, and third-born children spoke their parents' L1 even less often than did second-born children. We predicted that children who were NOT first- or only-born would be exposed to more English and might have larger receptive–expressive gaps than the children who were first- or only-born.

Methods

Participants

Kindergarten children were recruited during registration at two public elementary schools participating in a pre-reading program. Each school was located in the same neighborhood in Memphis, Tennessee. There was a large Hispanic population in the schools' neighborhood, and Hispanic children constituted a significant portion of the student body in each of the schools.

The present study included a bilingual group and a monolingual comparison group. Students were termed "bilingual" if caretakers reported Spanish spoken in the home. The bilingual group included 222 Hispanic K children, and the monolingual English-speaking group included 133 children (30 of these bilingual and 14 of these monolingual students were the research participants in Oller, Jarmulowicz, Gibson & Hoff, 2007). Since an important focus of the current study was the relationship of the receptive–expressive gap to demographic variables, we selected those participants for whom we had information regarding ALL of the demographic variables of interest. This reduced the bilingual group to 127. In order to limit the possibility that any single child's scores might inordinately affect the overall mean, we adopted a conservative approach and omitted outliers beyond three standard deviations from the mean in either direction on any of the vocabulary tests. This reduced the bilingual group for the primary analyses reported here to 124 (53 girls, 71 boys). The children's ages ranged from 5;0 to 7;1 (Mean = 5;7; SD = 5 months) at the time of entrance to the program in the fall school semester.

The English-speaking monolingual children participated in the same pre-reading program as the bilinguals. We included only those monolingual children for whom we had information across all demographic variables of interest, thus reducing the monolingual group to 110 (51 girls, 59 boys). Children's ages in the monolingual group ranged from 5;0 to 7;0 (Mean = 5;8; SD = 5 months).

Materials

Questionnaire

At the beginning of the study, when caregivers gave their informed consent, they filled out a questionnaire with the assistance of bilingual assistants if needed. Parents who did not complete the questionnaires were often contacted by a native Spanish-speaking assistant by telephone. In addition to traditional demographic information such as the child's age and parents' education, this questionnaire included information related to the nine predictions indicated above, including the child's English language background (child's age of English language exposure, attendance in English language preschool), child's country of birth, and household make-up (number of adults and children in the home and their relationships to the subject). We also requested information about the parents, specifically, their country of birth, length of residency in the USA, and English proficiency. Parents rated their English proficiency using a three-point scale, where 1 was no English proficiency, 2 was some English proficiency, and 3 was English proficient. The Spanish-speaking parents tended to rate themselves as having no English proficiency (84% rated themselves 1), a perhaps surprising fact given that their mean length of residency was 7.85 years. The questionnaire was available in both English and Spanish.

According to the questionnaire results, mothers of the bilingual participants in this study had low educational attainment (Mean = 8.13 years of formal education; SD = 3.05 years), had resided in the USA for a significant period of time (Mean = 7.8 years; SD = 4 years), spoke primarily Spanish to their children, and rated themselves as limited in English proficiency (1.19 on a three-point scale where 1 was no English proficiency, 2 was some English proficiency, and 3 was English proficient). All mothers reported that they were born in a Spanish-speaking country (91% in Mexico). Forty-five percent of the bilingual children had attended a preschool where

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

Table 1. *Means and (in parentheses) standard deviations for demographic variables.*

	Bilinguals (N = 124)	Monolinguals (N = 110)
Age average in months	67.22 (5.06)	68.41 (5)
Sex		
boys	57%	54%
girls	43%	46%
Birth country		
USA	73%	100%
other	27%	
Preschool in English		
yes	45%	51%
no	55%	49%
First English exposure average age in years	3.39(1.62)	N/A
Mothers' stay in USA average in years	7.85 (4)	N/A
Mothers' English proficiency		
none	84%	N/A
some	16%	N/A
Mothers' education average in years	8.13 (3.05)	11.98 (2.43)
Adults in home average	1.92 (1.16)	1.98 (.95)
Children in home average	3 (1.2)	2.51 (1.28)
Birth order		
first- or only-born	31%	N/A
later-born	69%	N/A

English was the language of instruction. Mothers of the monolingual English-speaking students had a mean of 11.98 (SD = 2.43) years of formal education and spoke only English. Fifty-one percent of the monolingual students had attended preschool. See Table 1 for a summary of the demographic information.

Standardized vocabulary testing

Standardized tests used in this investigation included the *Woodcock Language Proficiency Battery - Revised* (WLPB-R; Woodcock, 1991), its Spanish language equivalent the *Woodcock Language Proficiency Battery – Revised: Spanish form* (WLPB–RS; Woodcock & Muñoz-Sandoval, 1995), the *Peabody Picture Vocabulary Test – Third Edition* (PPVT–III; Dunn & Dunn, 1997), and its Spanish language equivalent the *Test de Vocabulario en Imagenes Peabody* (TVIP; Dunn, Padilla, Lugo & Dunn, 1986). All tests used in the current study have been widely used in both monolingual and bilingual language research. We chose the most widely respected receptive and expressive vocabulary tests with norms in both languages. The receptive and expressive tests were not produced by the same company, and consequently the norming groups were different. However, both Spanish language tests were normed on monolingual speakers ensuring that the standards of comparison for the Memphis results would be monolingual. Data from examiners' manuals provide strong indications of reliability and validity. Each test provides a mean standard score of 100 and a standard deviation of 15, making comparisons between tests straightforward.

The entirety of the PPVT–III and TVIP were administered, but only the Picture Vocabulary subtest of the WLPB–R was administered. This subtest provides a stand-alone score that can be compared to other stand-alone scores within the WLPB–R and to other language scores like the PPVT–III and TVIP.

Expressive vocabulary testing

Testers administered the picture vocabulary subtests of the WLPB–RS and WLPB-R. During the picture vocabulary subtest of the WLPB–RS, children provided the Spanish word associated with a realistic, color painting. The test allowed for some variation in Spanish dialect, for example, the word *perrito* "puppy" was accepted for the word *perro* "dog". A ceiling was reached when the child missed six consecutive items. The WLPB-R followed the same procedures and rules as its Spanish counterpart.

The WLPB-R examiner's manual provided the norming procedures involved in the development of the test. Participants in the norming sample matched 1980 US Census data across 10 variables, including census region,

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

community size, gender, race, Hispanic/not Hispanic, public/private university, two-year/four-year university, adult education level, adult employed/unemployed, and adult type of job (white collar/blue collar/service). Therefore, the norming sample closely approximated the demographic makeup of the USA. The design of the picture vocabulary subtest parallels real-life language performance requirements and makes it unlikely that there are confounding variables in the results, which supports its validity as a measure of productive vocabulary.[1]

The Spanish-speaking norms were developed based on the scores from approximately 2,000 native Spanish speakers residing in Costa Rica, Mexico, Peru, Puerto Rico, Spain, and the United States. A survey of the USA participants, which made up 34% of the sample, showed that they closely approximated monolingual Spanish speakers in terms of extent of Spanish language use.

Receptive vocabulary testing

On the PPVT–III, testers uttered an English word, and the child pointed to the corresponding item from a field of four line-drawings arranged in a two-by-two configuration. A ceiling was reached when the child missed 8 from any set of 12 consecutive items. The standardized average is 100 with a standard deviation of 15. Two forms (A and B) were counterbalanced across pre- and post-test. Each form contained 175 items.

The original PPVT was published in 1959 (Dunn, 1959) and revised in 1981 (PPVT–R; Dunn & Dunn, 1981). The PPVT–III was the 1997 extension to this well-established test, which was updated to reduce cultural biases. The test was normed on monolingual English-speaking residents in the USA whose demographic qualities were matched to 1994 US Census data. These demographic qualities included age, race, gender, education level, region of residence, and population density. As reported in the examiner's manual, 20 categories of nouns, descriptors, and verbs were represented in the test. The test makers claim that the test covers most of the ideas that children come into contact with; however, they concede that some ideas are not amenable to illustration with pictures. The examiner's manual contains an in-depth explanation of the process of word item selection and standardization, which supports its validity as a measure of monolingual receptive English vocabulary.

The TVIP (Dunn et al., 1986) is the Spanish adaptation of the PPVT–R (Dunn & Dunn, 1981). Administration and physical orientation of the test materials are identical to the PPVT–III (Dunn & Dunn, 1997). Test takers point to an uttered item that corresponds to one of four pictures. A ceiling is reached when the participant misses six of eight consecutive items. The mean is 100 and the standard deviation is 15.

To develop the TVIP, items from the two forms of the PPVT–R were translated into Spanish (350 items), and these items were administered to two monolingual Spanish-speaking groups. One group was located in Mexico City, the other in Puerto Rico. Socioeconomic data were not obtained for the Mexico City group but were obtained for the Puerto Rico group. Test makers attempted to match participant demographics to the Census data of Puerto Rico. For age groups that were not matched to Census data, scores were weighted to statistically match Census data. The scores from the Mexico City and Puerto Rico groups were combined, and items were calibrated for difficulty. Ultimately, 125 items from the PPVT–R's 350 items were used for the TVIP, ordered by difficulty based on actual test scores from the norming groups.

Procedure

Testers

A clinically certified speech-language pathologist trained the testers in the administration of the standardized vocabulary tests. Training included practice testing with adults and children. In the first year of this five-year study, non-native but fluent Spanish speakers administered both the Spanish and English vocabulary tests. In subsequent years, the Spanish tests were given by native Spanish speakers.

Test administration

At the beginning of the school year, participants were given a battery of language tests that included the vocabulary tests described above. Testing was conducted in the school at times convenient for the teachers. Effort was made to test in quiet spaces with minimal distractions. Monolingual English-speaking children took only English tests, while bilingual children took tests in both English and Spanish. All tests were administered and scored according to published guidelines. We systematically balanced the administration between Spanish and English testing, with half of the participants receiving the Spanish testing first and half receiving English testing first.

Data analyses

Typically developing monolinguals should not exhibit a receptive–expressive gap (i.e., on average, children taking a receptive vocabulary test should obtain standard scores more or less equivalent to those they obtain on an expressive vocabulary test in the same language). Our research focused on contrasting the performance of the monolingual and bilingual children in the current study. The nine demographic variables from the questionnaire

[1] In spite of its widespread usage, methodological concerns have been expressed about the WLPB–R expressive vocabulary test in Spanish. Oller et al. (2010) offer perspective on the concerns.

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

Table 2. *Means and (in parentheses) standard deviations for vocabulary testing.*

	PPVT-III	WLPB-R	TVIP	WLPB–RS
Bilinguals (124)	60.86 (17.74)	53.53 (20.77)	88.23 (17.1)	67.21 (20.6)
Monolinguals (110)	89.84 (14.10)	96.32 (18.72)		

Table 3. *Average scores and effect sizes (Cohen's* d*) for the bilingual children's English vocabulary tests according to whether the gap was predicted to be large or small.*

	English testing							
	Gap predicted to be small	RecEng	SD	ExpEng	SD	Gap	N	d
High English Exposure	USAborn	63.26	17.29	56.16	20.07	7.1	91	0.38
	EngPresch	63.80	18.06	58.77	21.38	5.03	56	0.25
	EarlyEngExposure	60.60	19.46	56.09	21.34	4.51	53	0.22
	MothResLong	62.12	17.99	55.31	20.29	6.81	107	0.36
	MothEngHigh	70.60	19.68	63.65	21.33	6.95	20	0.34
	MothEdHigh	61.30	18.35	55.26	22.13	6.04	76	0.30
	FewAdultsInHome	61.28	18.72	54.73	21.97	6.55	96	0.32
	ManyChildrenHome	62.32	17.50	54.87	18.84	7.45	77	0.41
	SecondOrLaterBorn	60.78	17.79	54.15	19.62	6.63	86	0.35
	Gap Predicted to be large	RecEng	SD	ExpEng	SD	Gap	N	d
Low English Exposure	NotUSAborn	54.24	17.55	46.27	21.22	7.97	33	0.41
	NotEngPresch	58.44	17.23	49.22	19.36	9.22	68	0.50
	LateEngExposure	61.06	16.49	54.79	51.62	6.27	71	0.16
	MothResShort	52.94	14.13	42.35	20.82	10.59	17	0.60
	MothEngLow	58.99	16.81	51.59	20.19	7.4	104	0.40
	MothEdLow	60.17	16.90	50.79	18.3	9.38	48	0.53
	MoreAdultsInHome	59.43	14.10	49.43	15.64	10	28	0.67
	FewChildrenHome	58.47	18.07	51.34	23.64	7.13	47	0.34
	FirstOrOnlyBorn	60.61	17.87	52.13	23.39	8.48	38	0.41

RecEng = Receptive English mean standard score; ExpEng = Expressive English mean standard score; Gap = RecEng minus ExpEng

that we reasoned might be related to degree of exposure to English and thus to the magnitude of a potential receptive–expressive gap were termed: (i) birth country, (ii) English preschool attendance, (iii) age of English exposure, (iv) mother's length of USA residency, (v) mother's English proficiency, (vi) mother's education, (vii) number of adults in the home, (viii) number of children in the home, and (ix) birth order (children who were first-born or without siblings vs. others). In order to make group comparisons, we changed continuous variables into dichotomous variables corresponding to presumed high and low English exposure.

We calculated means and standard deviations of the standardized vocabulary tests and compared the receptive–expressive gap in the bilingual children in L1 across the nine selected demographic variables. The size of the receptive–expressive gap and its Cohen's *d* effect sizes were calculated for each variable. We then conducted ordinary least squares regression analyses to determine the extent to which the demographic variables reported above contributed first to prediction of the receptive and expressive vocabulary score outcomes and second to the magnitude of the receptive–expressive gap. In order to capture as much statistical power as possible, continuous variables were not converted to dichotomous variables for the regression analysis.

Results

For bilingual students, the overall average of the standardized expressive vocabulary scores in Spanish (L1) was 67.21 (20.6), whereas the receptive average was

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

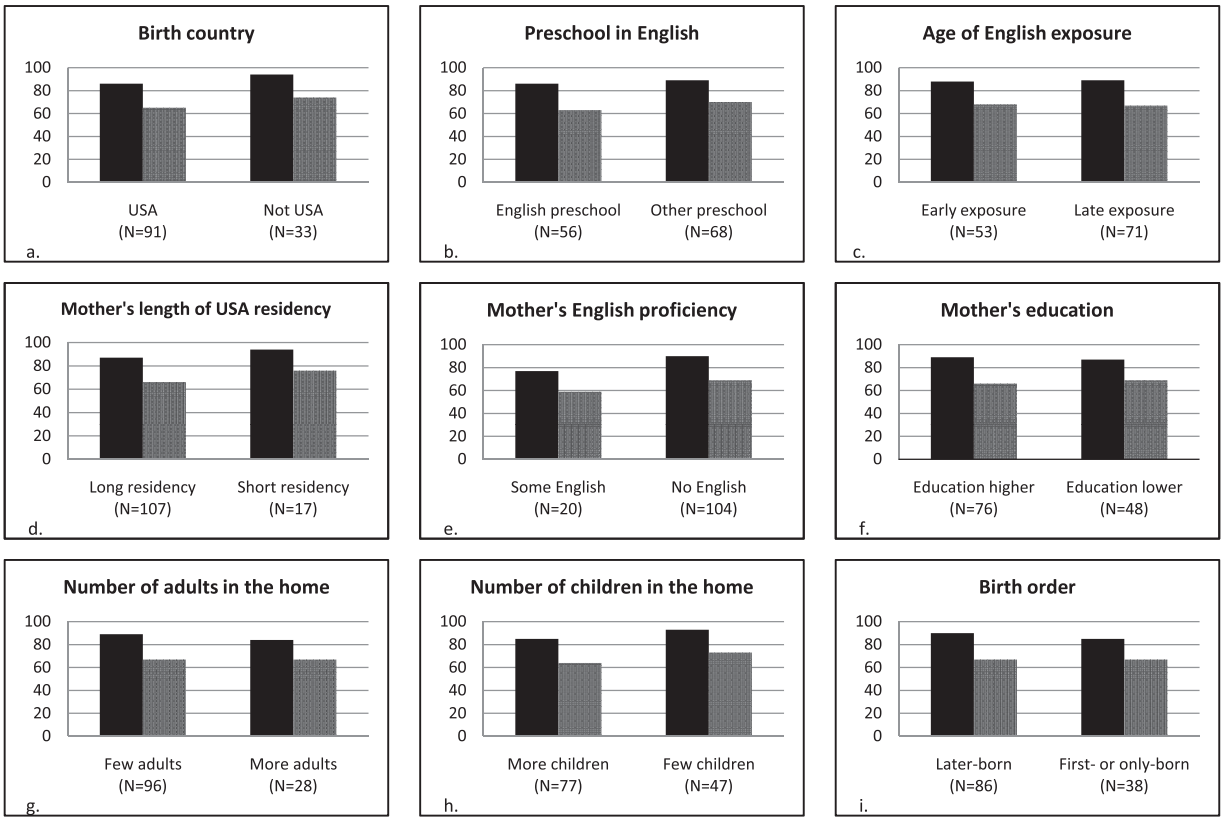

Figure 1. Histograms of children's receptive and expressive vocabulary performance in Spanish relative to the nine dichotomous variables: (a) birth country, (b) preschool attendance, (c) age of English exposure, (d) mother's length of USA residency, (e) mother's English proficiency, (f) mother's education level, (g) number of adults in the home, (h) number of children in the home, and (i) birth order. Black represents receptive vocabulary; gray represents expressive vocabulary.

88.23 (17.1), indicating an average receptive–expressive gap of 21.02 points, a difference of nearly one-and-a-half standard deviations. These same children scored 53.53 (20.77) on the English expressive vocabulary test and 60.86 (17.74) on the English receptive vocabulary test, a receptive–expressive gap of only 7.33 points. A MANOVA revealed that the interaction between test language and the difference between receptive and expressive vocabulary was also highly significant ($p < .0001$), indicating that the degree of the receptive–expressive gap was much greater in Spanish than in English for the bilingual children, and the effect size of the difference between the gap in the two languages was large (Cohen's $d = 0.87$). The difference between receptive and expressive scores was highly significant for both languages ($p < .0001$), although in the case of L1 (Spanish) the effect size (Cohen's $d = 1.11$) was large, according to standard rules of thumb, and in L2 the effect size ($d = 0.38$) was small.

Monolingual English-speaking children in the comparison group (N = 110) showed a six-point discrepancy in their expressive and receptive vocabulary scores, and this discrepancy lay in the OPPOSITE DIRECTION of the one found in the Spanish vocabulary scores. The average standardized receptive vocabulary score was 89.84 (14.10) and the average expressive vocabulary score was 96.32 (18.72). The difference was highly significant by t-test ($p < .0001$), but the effect size was small ($d = 0.39$) (see Table 2).

We predicted that nine variables would influence the bilingual children's English test scores, because the variables reflected differences in English exposure. The data showed a pattern suggesting that the nine variables as a group predicted English skills in the bilingual children, and thus can be interpreted as predictors of degree of English exposure (see Table 3). For eight of the nine demographic variables, English receptive scores were higher for the high English exposure groups than for the low English exposure groups (mean difference = 4.6 standard points). Likewise, for all nine of the variables, English expressive scores were higher for the high English exposure group (mean difference = 6.79 standard points) than for the low English exposure group. This tendency for higher English performance in the high English exposure groups for 17 out of 18 comparisons (2 degrees of English exposure × 9 demographic variables) was statistically

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

Table 4. *Average scores and effect sizes (Cohen's* d*) for the Spanish vocabulary tests according to whether the gap was predicted to be large or small.*

	Spanish testing							
	Gap predicted to be large	RecSpan	SD	ExpSpan	SD	Gap	N	d
High English Exposure	USAborn	86.05	17.19	64.73	18.72	21.32	91	1.19
	EngPresch	86.23	16.48	63.38	17.33	22.85	56	1.35
	EarlyEngExposure	87.72	17.77	67.60	19.76	20.12	53	1.07
	MothResLong	87.29	17.12	65.84	19.95	21.45	107	1.15
	MothEngHigh	77.10	17.29	58.95	18.05	18.15	20	1.03
	MothEdHigh	88.78	16.20	65.78	19.59	23.00	76	1.28
	FewAdultsInHome	89.34	17.32	67.28	21.06	22.06	96	1.14
	ManyChildrenHome	85.35	16.52	63.92	19.08	21.43	77	1.20
	SecondOrLaterBorn	89.86	15.82	67.10	20.59	22.76	86	1.24
	Gap Predicted to be small	RecSpan	SD	ExpSpan	SD	Gap	N	d
Low English Exposure	NotUSAborn	94.24	15.57	74.06	24.04	20.18	33	1.00
	NotEngPresch	89.88	17.54	70.37	22.57	19.51	68	0.97
	LateEngExposure	88.62	16.70	66.92	21.32	21.70	71	1.13
	MothResShort	94.18	16.23	75.82	23.06	18.36	17	0.92
	MothEngLow	90.37	16.29	68.80	20.75	21.57	104	1.16
	MothEdLow	87.38	18.59	69.48	22.11	17.9	48	0.88
	MoreAdultsInHome	84.42	16.02	66.96	19.24	17.46	28	1.00
	FewChildrenHome	92.96	17.15	72.60	22.02	20.36	47	1.03
	FirstOrOnlyBorn	84.55	19.41	67.45	20.76	17.10	38	0.85

RecSpan = Receptive Spanish mean standard score; ExpSpan = Expressive Spanish mean standard score; Gap = RecSpan minus ExpSpan

significant by a sign test ($p = .0002$). The exception was that lower receptive scores were associated with the EARLY ENGLISH EXPOSURE variable, than with the LATE ENGLISH EXPOSURE variable.

In contrast, our prediction that those children with higher English exposure would exhibit larger receptive–expressive gaps in Spanish than children with lower English exposure was not generally borne out in the results. Figure 1 shows histograms for the nine variables with respect to the receptive–expressive gap. Seven of the nine variables showed a slightly higher receptive–expressive gap for the higher English exposure group (see Table 4); however, the early English exposure group and the children whose mothers spoke English did not follow this pattern. The mean difference in gap sizes across the nine was only 2.11 points, and none of these nine differences in gap size was significantly different by one-way ANOVA with Bonferonni correction.

An intriguing finding is suggested by the general failure of the predictions about the receptive–expressive gap in Spanish. Children in both the high and the low English exposure groups showed ROBUST receptive–expressive gaps in Spanish across all nine variables. Table 4 presents the means and standard deviations for the receptive and expressive tests along with the receptive–expressive gap and the effect sizes for the differences between the receptive and expressive test scores for each group.[2] Not only were the receptive scores in Spanish statistically significantly higher than expressive scores in EVERY ONE of the 18 possible comparisons ($p < .001$), but also every one of the 18 effect sizes was large (obtained range $d = 0.85$–1.35).

Given that the receptive–expressive gap in Spanish (L1) vocabulary was extremely robust across all the variables, we performed ordinary least squares regression analyses on data from the 124 children for whom we had all the demographic data to test effects of the nine selected demographic variables and to determine if any of those variables uniquely contributed to the magnitude of effects on vocabulary skills. In order to provide perspective regarding the effect of the demographic variables on

[2] The pattern of results in Table 4 did not change when we applied the analysis to all of the children in the original study, although not every one of those participants had data for every variable. In addition, we analyzed age (5;7 and younger vs. older than 5;7) and gender and found no significant differences in the gap for age or gender in Spanish; however, girls did better than boys on the receptive vocabulary test in Spanish $F(1,122) = 8.43$; $p = .004$).

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

Table 5. *Summary of regression analysis for values predicting vocabulary scores in Spanish.*

Variable	B	Standard Error	Beta
Birth country	−10.84	3.86	−.28*
English preschool	−6.28	2.90	−.19*
Age of English exposure	−1.17	.97	−.11
Mother's length of USA residency	.33	.44	.079
Mother's English proficiency	–7.97	3.44	−.21*
Mothers' education	.28	.47	.05
# Adults in home	−.66	1.32	−.05
# Children in home	−4.06	1.26	−.29*
Birth order	3.58	3.34	.10

$F(9,114) = 3.27$; $p = .001$; $R^2 = .21$; * $p < .05$

Table 6. *Summary of regression analysis for values predicting vocabulary scores in English.*

Variable	B	Standard Error	Beta
Birth country	46.89	4.12	.15
English preschool	7.96	3.09	.22*
Age of English exposure	1.66	1.04	.15
Mother's length of USA residency	.79	.47	.18
Mother's English proficiency	6.02	3.67	.15
Mothers' education	1.05	.50	.18*
# Adults in home	−.91	1.41	−.06
# Children in home	.85	1.35	.06
Birth order	1.60	3.56	.04

$F(9,114) = 3.49$; p = .001; $R^2 = .22$; * $p < .05$

vocabulary performance in general, we first used overall Spanish vocabulary as the dependent variable, which was determined by averaging the Spanish receptive and expressive vocabulary scores. We entered all variables into the model as a single entry. The overall model was significant, $F(9,114) = 3.27$; $p = .001$, and accounted for 21% of the variance in the overall Spanish vocabulary scores. Four of the nine variables uniquely contributed to the variance explained: The child's country of birth ($\beta = –0.28$), attendance in English language preschool ($\beta = –.19$), the mother's self-rated English proficiency on a three-point scale ($\beta = –.21$), and the number of children (including the participating child) living in the home ($\beta = –.29$) (see Table 5).

Table 7. *Summary of regression analysis for values predicting the receptive–expressive gap in Spanish.*

Variable	B	Standard Error	Beta
Birth country	−52.00	4.27	−.01
English preschool	3.54	3.20	.10
Age of English exposure	.60	1.08	.06
Mother's length of USA residency	.66	.49	.15
Mother's English proficiency	−4.55	3.80	−.12
Mothers' education	1.06	.52	.19
# Adults in home	.40	1.46	.03
# Children in home	.70	1.40	.05
Birth order	4.09	3.70	.11

$F(9,114) = 1.20$; $p = .30$

We performed the same analysis using the overall English vocabulary score from the bilingual children as the dependent variable. This model was also significant $F(9,114) = 3.49$; $p = .001$, explaining 22% of the variance in the overall English vocabulary score. Only two of the nine variables, attendance in an English-speaking preschool ($\beta = .22$) and mother's education level ($\beta = .18$), uniquely contributed to the variance explained (see Table 6).

On the other hand, the same overall model for the receptive–expressive gap in Spanish as the dependent variable was not statistically significant, $F(9,114) = 1.20$, $p = .30$ (see Table 7). Therefore, these variables did not predict the size of the gap between the Spanish vocabulary scores. The receptive–expressive gap in Spanish remained relatively constant regardless of the demographic variable, and despite the influence of these demographic variables on the test scores themselves.

Discussion

For the children in this study, there was a gap between receptive and expressive L1 vocabulary (Spanish) that was very strong and reliable in every subgroup comparison that was made. We predicted that children with more exposure to English might present with larger receptive–expressive gaps than children with less exposure to English based on the fact that amount of language input correlates positively with larger vocabularies (Hart & Risley, 1995; Hoff, 2006) and that bilingual children tend to have larger vocabularies in the language in which they have more exposure (Patterson, 2002).

While our prediction failed to materialize, we are struck by the robustness of the receptive–expressive gap, which showed large effect sizes in all the comparisons of children from high and low English-exposure environments. Spanish (L1) receptive vocabulary was

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

on average 21.02 points higher than L1 expressive vocabulary, which represented a difference of 1.40 standard deviations. When these same children were tested in English (L2), a much smaller receptive–expressive gap, of 7.33 points, was observed, representing 0.49 of a standard deviation. Because the scores in the current study were standardized, they cannot be dismissed as merely reflecting the inherently greater cognitive load for expressive as opposed to receptive language tasks. Instead, these standardized scores appear to reflect a real difficulty these children had when accessing their productive L1 vocabularies in the circumstance of L2 immersion. The regression model that included all nine demographic variables under consideration was statistically significant when applied to the average Spanish or English vocabulary scores (accounting for 21% and 22% of variance, respectively), but was statistically nonsignificant when applied to the receptive–expressive gap itself.

Two plausible conclusions can be drawn from these results. First, children in this study had limited expressive access to their Spanish (L1) vocabularies, evinced by receptive vocabulary standard scores that were much higher than expressive vocabulary standard scores. When these same children were tested in English (L2), their receptive vocabulary standard scores were much less discrepant than their expressive vocabulary standard scores, suggesting that the L1 receptive–expressive gap is a reflection of a specific loss of expressive access to L1 vocabulary in the L2 immersion circumstance, where it appears that L1 activation is particularly low.

The second, more tentative, conclusion is that the onset of the receptive–expressive gap occurred abruptly, because even children whose parents reported they had started learning English only with entry to K showed a robust receptive–expressive gap in Spanish. This conclusion is tempered by the possibility that they had an L1 receptive–expressive gap before they arrived at K. As Oller, Jarmulowicz, Pearson & Cobo-Lewis (2010) have noted, if the total amount of linguistic input that the child had experienced were the primary factor responsible for the receptive–expressive gap, one would expect the L1 expressive vocabulary deficit to reveal itself slowly, as the cumulative result of increasing input from English. However, in the present study, even children who were born outside of the USA or who had not attended preschool in English showed the receptive–expressive gap within the first months of immersion in English-only K.

Given its robustness and magnitude, its ostensible abruptness, as well as the likelihood that low expressive vocabulary performance in L1 might be interpreted as a language-learning deficit, it is important to expatiate on possible explanations for the receptive–expressive gap in L1.

Possible explanations

A suppression effect

How do speakers put words to the ideas they want to express? Researchers have created a variety of models to answer this question. In many adult models of monolingual speech production, a target concept is generated, a word that expresses the concept is selected, the phonological features of the word are organized, and instructions are sent to the muscles involved for the production of those features (Levelt, 1989; Levelt, Roelofs & Meyer, 1999). In some models, the process is complicated for bilingual speakers, because they access words from potentially two languages. Some models, such as Green's (1998) Inhibitory Control model, as well as others (Hermans, Bongaerts, De Bot & Schreuder, 1998; Lee & Williams, 2001), have invoked top–down, inhibitory control as the means by which bilingual lexical selection takes place.[3] In these models, a concept activates the shared lexico-semantic field, and words from both languages compete for selection. Once the competitors are activated, a top–down mechanism suppresses activation of words from the non-target language. This allows a word from the target language to be selected. For example, in a Spanish–English bilingual the concept "table" activates both the English word *table* and the Spanish word *mesa*. If the target language is Spanish, the activation of *table* is suppressed, which means that *mesa* is preferred for selection. Essential to this model is the idea that the suppression is reactive: Words are first activated in both languages and then inappropriate ones are suppressed on each occasion of lexical access.

We cannot rule out the possibility that children in the present study experienced a sort of inhibitory control (which we will call a suppression effect) on Spanish; however, the data suggest some modification to Green's (1998) Inhibitory Control model. Our thinking is in accord with the proposal by Linck, Kroll & Sunderman (2009) that L1 becomes inhibited in the context of L2 immersion. Linck et al. (2009) compared adults who studied a foreign language (L2) in a study-abroad program to adults who studied a foreign language in a traditional classroom in the L1 country. Participants in the study-abroad program showed reduced verbal fluency skills in L1, while the comparison group did not. In the present study, children presented with difficulty accessing L1 in an L2 context, a pattern resembling the study-abroad model.

[3] For the bilingual speaker, it has been assumed that both languages share the same conceptual system (Kroll & Stewart, 1994). In inhibitory models, the two languages also share the same lexico-semantic space. Costa (2005) pointed out that this is the view held by most researchers in psycholinguistics, even those who do not support an inhibitory control mechanism. This view seems to be supported by neuroscience, since the bilingual's languages appear to occupy the same neural tissue (see Stowe & Sabourin, 2005).

https://doi.org/10.1017/S1366728910000490 Published online by Cambridge University Press

This suggests a more proactive mechanism than the one proposed in the Inhibitory Control model (Green, 1998). Our data suggest the possibility that the L1 (Spanish) is preemptively suppressed in the circumstance of school-based L2 (English) immersion BEFORE the occasion of any single lexical access event. This makes it difficult to access the lexicon of the language that is not part of the current context.

There is actually another possible interpretation of the present outcomes that does not require a specific suppression mechanism. It is conceivable that what changes at the beginning of L2 immersion is a state of relative ACTIVATION of each language. As English becomes highly activated, Spanish may passively become relatively deactivated. Some researchers of lexical access and adult bilingualism have specifically argued against a suppression mechanism and in favor of a more conservative relative activation approach (Costa, Santesteban & Ivanova, 2006).

Whatever the mechanism is, suppression or relative deactivation, it appears to change relative to a bilingual's level of proficiency. This was suggested in results of Costa & Santesteban (2004) who compared the amount of time it took bilingual adults to switch between languages when naming a series of pictures. A delay at the moment when the target language switched indicated two things. First, the non-target language had been suppressed while the other language had been spoken. Second, at the moment of the language switch, the suppression of the to-be-accessed language persisted even as the speaker tried to overcome the suppression. Longer delays at accessing the previously suppressed language indicated stronger levels of suppression. Costa & Santesteban (2004) found that low L2 proficiency speakers had longer delays switching from L2 to the dominant L1 than vice versa, but no such asymmetry existed for speakers who were proficient in both L2 and L1. This suggests that the effects of suppression or relative activation are more salient when beginning to learn a new language. This line of reasoning is supported by the results of the Miami project (Oller & Eilers, 2002). Spanish–English bilingual children in the Miami study showed a reduction in the strength of the receptive–expressive L1 gap over time (at K, 28 standardized points; at fifth grade, 19 points).

The suppression (or relative deactivation) effect could be an effective, natural mechanism to help the L2 learner acquire the new language, and its power could be thought to diminish over time as command of the second language grows and the speaker has less need to block interference from the first language or more capability to quickly raise the activation level of either language. Thus, particularly at the beginning of L2 learning, the receptive–expressive gap may be most apparent, because that is when the learner needs most to keep L1 out of the way, and may be least capable of quickly raising activation levels. We would expect, then, that longitudinal data might show a reduction of the gap with time, which would reflect increased ease of access to the L1 in the L2 context.

The role of tasks in the receptive–expressive gap

It is not disputed that, in general, receptive vocabulary tasks require less effort than production tasks. When children are tested for receptive language skills on a test like the TVIP used in the current study, they are presented with four pictures and asked to point to the picture that matches the word uttered by the tester. When they are tested for expressive skills they are asked to name pictures. Oller et al. (2010) reasoned that whatever the differences are between effort level required in these two tasks, the discrepancy between them may be magnified in a circumstance of low language activation (or suppression). We might say that this implies that the tests of receptive and expressive vocabulary are normed for a circumstance (high activation of the language in question) that does not apply in the case of L2 immersion for new language learners. When activation for a language is low (or suppression is high), the greater effort required for picture naming may be intensified with respect to picture pointing. This reasoning might help explain why there was both a large receptive–expressive gap for L1 and also a smaller one for L2. The smaller gap for L2 might suggest that activation of English, while higher than for Spanish in the English immersion circumstance, was still lower than it would be for monolinguals.

The social domain: Peer effects

Another possible contributor to the receptive–expressive gap relates to how the child is situated in the peer culture. Corsaro & Eder (1990) in their ethnographic study of preschool children asserted that children play an active role in the development of their own peer culture. They indicate that "[c]hildren make persistent attempts to gain control of their lives and to share that control with each other" (p. 202). This control is manifested in the development of a peer culture whose members are preoccupied by social participation. After gaining access to the peer culture, children protect their position in it and, in so doing, create barriers to new entrants. Peers who want to join in the group must overcome that barrier by developing "complex access strategies" (p. 203). These sometimes language-dependent strategies are part of the social skills the child develops to participate with peers. Children who speak a minority language have been shown to have problems with peer group access in much the same way as do children with language learning difficulties (Gertner, Rice & Hadley, 1994).

The desire to belong to a peer group appears to be universal. Baumeister & Leary (1995, p. 522) wrote that "it seems fair to conclude that human beings are

fundamentally and pervasively motivated by a need to belong". This drive can be so strong as to lead to self-harm if one is rejected by his peer group (Gilbert, 2003). If the children in the current study were motivated to belong to the peer culture, and if language plays a role in acceptance by that peer culture, then the bilingual children in the current study may have perceived themselves as being at a social disadvantage when compared to monolingual English-speaking children due to their lack of English knowledge. Furthermore, they may have felt themselves to be at risk of social rejection and might have been willing to go to some lengths to avoid it.

Speaking Spanish is a signal that the child is not a part of the English-speaking peer culture. This might motivate the child, either consciously or unconsciously, to avoid the use of Spanish in an English-speaking context, not for language learning purposes but to conceal that he/she is not a member of the group, thus avoiding negative social consequences. The reluctance to signal their Spanish-speaking status might have been reflected in a picture naming task in Spanish.

In our characterization of a suppression or differential activation mechanism, context is important. We have argued for the possibility that L1 becomes inhibited in the context of school-based L2 (English) immersion. Presumably, a change in context could precipitate a change in the salience of the suppression (or differential activation) mechanism on L1. If the recognition and naming tests were administered in the homes or in the birth countries of the children in the current study (during vacation trips, for example), one might find that the receptive–expressive gap disappears, because English would not be used in these contexts, and there would be little motivation to suppress Spanish L1. This is a testable empirical question for future research.

Clinical implications

The identification of the receptive–expressive gap introduces a new clinical wrinkle into the assessment of language disorders in bilingual children. Previous research (Oller, Pearson & Cobo-Lewis, 2007; Pearson, Fernández & Oller, 1995; Umbel & Oller, 1994; Umbel, Pearson, Fernández & Oller, 1992) has suggested that bilingual children's vocabularies are dependent on experiences. Some words, like *dedal*, the Spanish word for *thimble*, are likely only to be heard at home and unlikely to appear in the English environment (school). Consequently, the child may acquire the word in Spanish but not English.

The tendency to know words in L1 but not L2 and vice versa is what Oller, Pearson and Cobo-Lewis (2007, p. 192) termed the "distributed characteristic". Vocabulary is distributed across the two languages. If the child is tested in only one language, the result might inaccurately indicate an impoverished vocabulary, when, in fact, many of the words in question might be known in the untested language. Clinicians have been directed to test bilingual speakers in both of their languages to avoid false positives in testing (Caesar & Kohler, 2007). The assumption is that testing both languages will provide a more accurate measure of the child's true vocabulary knowledge.

The receptive–expressive gap, however, introduces a new concern. Whatever the underlying reasons for the phenomenon, confrontation naming tasks like the one in the current study are likely to indicate an impoverished L1 expressive vocabulary. Even L2 expressive vocabulary may be somewhat reduced relative to receptive vocabulary. This might in turn result in the over-identification of L2 learners as having expressive language deficits when in fact their skills are typical of the L2 learner profile. Possibly, testing children in other circumstances (at home or after a trip to an L1-speaking country) would show improved performance in picture naming in L1, although this remains untested. Kohnert (2004) proposed that the better indicator of language disorder in bilinguals is not standardized vocabulary tests, which are knowledge-based and experience-dependent, but instruments that assess processing. Dynamic assessment approaches (Lidz & Peña, 1996) that use test–teach–retest protocols might accomplish this goal, and help avoid false labeling of children as language-impaired, when in fact, they may be merely in the normal process of rapid language shift.

Limitations

There were some limitations to the current study's methodology. The demographic questionnaire asked mothers to rate their English proficiency on a scale of one to three, which may not have captured fine-grained distinctions between mothers. A Likert-style scale with a broader point range would have led to more variation and thus might have provided a more realistic picture of the proficiency levels of these mothers. Results on the three-point scale should thus be interpreted with caution.

A curious result was the relatively better performance by the monolinguals on the expressive vocabulary test. Although we systematically balanced the administration between Spanish and English testing, we did not balance administration of the individual tests within languages; therefore, we cannot rule out a test effect. However, we suspect that the participants in our study were simply different from the norming sample with respect to the balance of receptive and expressive skills. This could be a cultural or dialectal effect. On the other hand, it could be an effect of a school environment in which children in these schools received more encouragement to speak in the school environment than were children in the norming sample.

Finally, we urge caution in interpreting the size of the receptive–expressive gap in Spanish relative to the receptive–expressive gap in English for the L2 learners. The L2 receptive–expressive gap may have been influenced by floor effects, because of very low receptive and expressive vocabularies in English in the bilingual children.

Conclusion

The receptive–expressive gap was remarkably robust in this study. The gap has been alluded to in other studies, but rarely directly examined. We found evidence that Spanish–English bilingual children show low scores in L1 (Spanish) expressive vocabulary, but not in L1 receptive vocabulary. Furthermore, regardless of how exposure to English was measured, we did not find a reliable predictor of the L1 receptive–expressive gap. Even children with limited English exposure showed the receptive–expressive gap within three months of beginning K.

We conclude that the receptive–expressive gap is a perfectly natural phenomenon experienced by children immersed in a second language upon entering school and possibly in any L2 immersion environment. The results have important implications for models of bilingual lexical access and for clinical assessments of language problems. When generating models of lexical access, researchers should consider the possibility that a suppression mechanism (or relative deactivation) is operating on L1 in the L2 immersion circumstance. Furthermore, clinicians should consider the potential for the imbalances of skills related to the receptive–expressive gap to be falsely interpreted as a language disorder.

References

Baumeister, R. F., & Leary, M. R. (1995). The need to belong – desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117 (3),* 497–529.

Caesar, L. G., & Kohler, P. D. (2007). The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Language Speech and Hearing Services in Schools, 38 (3),* 190–200.

Corsaro, W. A., & Eder, D. (1990). Children's peer cultures. *Annual Review of Sociology, 16,* 197–220.

Costa, A. (2005). Speech production in bilinguals. In T. Bhatia & W. Ritchie (eds.), *The handbook of bilingualism*, pp. 201–223. Malden, MA & Oxford: Blackwell.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50 (4),* 491–511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32 (5),* 1057–1074.

Dunn, L. M. (1959). *Peabody Picture Vocabulary Test (PPVT)*. Circle Pines, MN: AGS Publishing.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test – Revised (PPVT–R)*. Circle Pines, MN: AGS Publishing.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test – Third Edition (PPVT–III)*. Circle Pines, MN: AGS Publishing.

Dunn, L. M., Padilla, E., Lugo, D., & Dunn, L. M. (1986). *Test de Vocabulario en Imagenes Peabody (TVIP)*. Circle Pines, MN: AGS Publishing.

Gertner, B. L., Rice, M. L., & Hadley, P. A. (1994). Influence of communicative competence on peer preferences in a preschool classroom. *Journal of Speech and Hearing Research, 37 (4),* 913–923.

Gilbert, P. (2003). Evolution, social roles, and the differences in shame and guilt. *Social Research, 70 (4),* 1205–1230.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1 (2),* 67–81.

Hakuta, K., & D'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics, 13 (1),* 72–99.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H Brookes.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition, 1 (3),* 213–229.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26 (1),* 55–88.

Hoff, E. (2009, April). Effects of dual language exposure on early lexical growth. Poster session presented at the Biennial Meeting of the Society for Research in Child Development, Denver, CO.

Kan, P. F., & Kohnert, K. (2005). Preschoolers learning Hmong and English: Lexical-semantic skills in L1 and L2. *Journal of Speech Language and Hearing Research, 48 (2),* 372–383.

Kohnert, K. (2004). Processing skills in early sequential bilinguals. In B. Goldstein (ed.), *Bilingual language development and disorders in Spanish–English speakers*, pp.53–76. Baltimore, MD: Paul H Brookes.

Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33 (2),* 149–174.

Lee, M., & Williams, J. N. (2001). Lexical access in spoken word production by bilinguals: Evidence from the semantic competitor priming paradigm. *Bilingualism: Language and Cognition, 4 (3),* 233–248.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). Multiple perspectives on word production. *Behavioral and Brain Sciences, 22 (1),* 61–75.

Lidz, C., & Peña, E. (1996). Dynamic assessment: The model, its relevance as a nonbiased approach, and its application to Latino American preschool children. *Language, Speech, and Hearing Services in Schools, 27 (4),* 367–372.

Linck, J., Kroll, J., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science, 20 (12),* 1507–1515.

Miccio, A., Tabors, P., Paez, M., Hammer, C., & Wagstaff, D. (2005). Vocabulary development in Spanish-speaking head start children of Puerto Rican descent. In J. Cohen, K McAlister, K Rolsted & J. MacSwan (eds.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism*, pp. 1614–1617. Somerville, MA: Cascadilla Press.

Morales, A., & Hanson, W. E. (2005). Language brokering: An integrative review of the literature. *Hispanic Journal of Behavioral Sciences, 27 (4),* 471–503.

Muñoz, M., & Marquardt, T. (2003). Picture naming and identification in bilingual speakers of Spanish and English with and without aphasia. *Aphasiology, 17 (12),* 1115–1132.

Oller, D. K., & Eilers, R. E. (eds.) (2002). *Language and literacy in bilingual children*. Clevedon: Multilingual Matters.

Oller, D. K., Jarmulowicz, L., Gibson, T., & Hoff, E. (2007). First language vocabulary loss in early bilinguals during language immersion: A possible role for suppression. In H. Caunt-Milton, S. Kulatilake & I. Woo (eds.), *Proceedings of the 31st Annual Boston University Conference on Language Development*, pp. 474–484. Somerville, MA: Cascadilla Press.

Oller, D. K., Jarmulowicz, L., Pearson, B. Z., & Cobo-Lewis, A. B. (2010). Rapid spoken language shift in early second language learning: The role of peers and effects on L1. In A. Durgunouglu & C. Goldenberg (eds.), *Dual language learners: Their development and assessment in oral and written language*, pp. 94–120. New York: Guilford Press.

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28 (2),* 191–230.

Patterson, J. (2002). Relationships of expressive vocabulary to frequency of reading and television experience among bilingual toddlers. *Applied Psycholinguistics, 23 (4),* 493–508.

Pearson, B. Z., Fernández, S., Lewedeg, V., & Oller, [D.] K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18 (1),* 41–58.

Pearson, B. Z., Fernández, S., & Oller, [D.] K. (1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language, 22 (2),* 345–368.

Shin, S. (2002). Birth order and the language experience of bilingual children. *TESOL Quarterly, 36 (1),* 103–113.

Stowe, L., & Sabourin, L. (2005). Imaging the processing of a second language: Effects of maturation and proficiency on the neural processes involved. *International Review of Applied Linguistics, 43 (4),* 329–353.

Swanson, H., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology, 46 (4),* 413–429.

Umbel, V. M., Oller, D. K. (1994). Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning, 44 (2),* 221–242.

Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development, 63 (4),* 1012–1020.

Windsor, J., & Kohnert, K. (2004). The search for common ground. Part I: Lexical performance by linguistically diverse learners. *Journal of Speech Language and Hearing Research, 47 (4),* 877–890.

Woodcock, R. (1991). *Woodcock Language Proficiency Battery – Revised (WLPB–R)*. Itasca, IL: Riverside Publishing.

Woodcock, R., & Muñoz-Sandoval, A. (1995). *Woodcock Language Proficiency Battery – Revised: Spanish Form (WLPB–RS)*. Itasca, IL: Riverside Publishing.

Yan, S., & Nicoladis, E. (2009). Finding le mot juste: Differences between bilingual and monolingual children's lexical access in comprehension and production. *Bilingualism: Language and Cognition, 12 (3),* 323–335.