Effect of Semantic Naming Treatment on Crosslinguistic Generalization in Bilingual Aphasia

Lisa A. Edmonds
Swathi Kiran
The University of Texas at Austin

Purpose: The effect of semantic naming treatment on crosslinguistic generalization was investigated in 3 participants with English–Spanish bilingual aphasia.
Method: A single-subject experimental designed was used. Participants received semantic treatment to improve naming of English or Spanish items, while generalization was tested to untrained semantically related items in the trained language and translations of the trained and untrained items in the untrained language.
Results: Results demonstrated a within- and across-languages effect on generalization related to premorbid language proficiencies. Participant 1 (P1; equal premorbid proficiency across languages) showed within-language generalization in the trained language (Spanish) as well as crosslinguistic generalization to the untrained language (English). Participant 2 (P2) and Participant (P3) were more proficient premorbidly in English. With treatment in English, P2 showed within-language generalization to semantically related items, but no crosslinguistic generalization. With treatment in Spanish, both P2 and P3 exhibited no within-language generalization, but crosslinguistic generalization to English (dominant language) occurred. Error analyses indicated an evolution of errors as a consequence of treatment.
Conclusions: These results are preliminary because all participants were not treated in both languages. However, the results suggest that training the less dominant language may be more beneficial in facilitating crosslinguistic generalization than training the more proficient language in an unbalanced bilingual individual.

KEY WORDS: crosslinguistic generalization, bilingual aphasia, naming treatment, language recovery

Bilingual aphasia is a loss of one or both languages in bilingual individuals whose language-dominant hemisphere (typically the left) has been damaged. Little is known about rehabilitation of bilingual aphasia, even though more than half of the world's population is bilingual, including growing bilingual populations in the United States (U.S. Bureau of the Census, 2000). To address the need for information regarding bilingual aphasia rehabilitation, the present study describes a theoretically motivated naming treatment designed to facilitate crosslinguistic generalization in English–Spanish bilingual individuals with aphasia.

Studies of bilingual aphasia have been largely limited to characterizing the nature of selective impairments (for a recent review, see Goral, Levy, & Obler, 2002). Thus far, treatment efforts have mainly involved individual case studies focused on broad language skills (Sasanuma & Suk Park, 1995; Watamori & Sasanuma, 1978) and have lacked detailed

1092-4388/06/4904-0729

pre–post treatment assessments in both languages (Junque, Vendrell, Vendrell-Brucet, & Tobena, 1989).

Results of the few studies that have systematically examined naming treatment in individuals with bilingual aphasia have been equivocal with respect to crosslinguistic generalization, the focus of the current study. For instance, cuing hierarchy treatment in English or Spanish did not yield crosslinguistic generalization for 1 Spanish–English bilingual with transcortical motor aphasia (TMA), as measured by the Bilingual Aphasia Test (BAT; Paradis, 1987) Naming subtest (Galvez & Hinckley, 2003; Hinckley, 2003). Hinckley used a cuing hierarchy with a balanced Spanish–English bilingual with TMA who was 4 months postonset. Language of treatment was alternated weekly. Overall improvement on the BAT (Paradis, 1987) was greater in Spanish than in English, but naming improvement on the BAT was equal across languages. Crosslinguistic generalization has been reported in cognates (e.g., *elephant* [*elefante*]) in 1 Spanish–English bilingual with TMA who received semantic and phonological treatment in both languages (Kohnert, 2004).

The proposed experiment attempted to address three important questions related to bilingual aphasia rehabilitation that previous studies have not answered and that have been raised in the literature (e.g., Costa, Santesteban, & Caño, 2005; Fabbro, 2001). First, is it sufficient to rehabilitate one language in patients with bilingual aphasia? Second, to what extent does rehabilitation in one language have beneficial effects in the untreated language? Third, to what extent does premorbid language proficiency affect recovery of each language? A theory-based approach modeled on previous naming studies with monolingual individuals with aphasia was integrated with recent models of bilingual memory to develop the current treatment protocol.

In monolingual aphasia, semantic treatments that are based on models of lexical processing and that focus on semantic features of items within a particular superordinate category (Drew & Thompson, 1999), or that attempt to facilitate spreading activation of semantically related words (semantic feature analysis), have been successful at facilitating phonological retrieval as well as generalization to untrained items (Boyle, 2004; Boyle & Coehlo, 1995; Lowell, Beeson, & Holland, 1995). These studies have been guided by established models of lexical access (Butterworth, 1989; Dell, 1986; Humphreys, Riddoch, & Quinlan, 1988; Levelt, Roelofs, & Meyer, 1999). These models assume that lexical access comprises two steps: lexical–semantic activation and phonological encoding. During lexical–semantic activation, semantic information for the target representation is accessed from the semantic system. During phonological encoding, target word forms are selected with regard to their phonological specification. Although most models of lexical access agree on these two steps, there is considerable debate regarding their temporal sequence (see Butterworth, 1989; Dell, 1986; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Humphreys et al., 1988; Levelt et al., 1999).

Current models of bilingual memory generally agree that bilingual individuals have a shared semantic/conceptual system with separate lexical representations of the two languages.[1] (Discussion of bilingual models will focus on points pertinent to the present study. See Kroll & de Groot, 2005, for an in-depth discussion of models.) However, the models differ on how the lexicons interact with the conceptual system and with each other, and these differences often depend on language proficiencies. For example, Kroll and Curley (1988) examined bilingual individuals with different proficiencies and found evidence for two different processing models. Evidence for the word association model (Potter, So, von Eckardt, & Feldman, 1984), which posits that second language words (L2) gain access to concepts only through first language mediation (L1), was found in low L2-proficiency bilinguals. In contrast, bilinguals with high proficiency in L1 and L2 show results in support of the concept mediation model. The concept mediation model (Potter et al., 1984) proposes that the second language lexicon directly accesses concepts and predicts that translation times from L1 to L2 and picture naming times in L2 should be similar because both require conceptual access prior to the retrieval of L2 lexical items. The revised hierarchical model (Kroll & Stewart, 1994) and the mixed model (de Groot, 1992) allow for language proficiency differences across bilingual individuals. The revised hierarchical model (Kroll & Stewart, 1994) proposes connections between L1 and L2 and between each lexicon and the central concept. The connections differ in their strengths as a function of fluency in L1 relative to L2. Lexical associations from L2 to L1 are assumed to be stronger than those from L1 to L2, and the links between the conceptual system and L1 are assumed to be stronger than from the conceptual system to L2. Like the revised hierarchical model, de Groot's (1992) mixed-model proposes connections that can differ in strength depending on relative proficiency of languages, including bilingual individuals proficient in both L1 and L2, for whom connections between the conceptual system and both lexicons and between each lexicon are equally strong. Thus, the mixed model is the most flexible model because it allows for a range of language proficiencies.

Another question concerning bilingual lexical access involves whether activation of the semantic system spreads to one or both lexicons during phonological

[1]In this paper, we use the terms *semantic system* and *conceptual system* interchangeably based on a recent discussion in the bilingualism literature regarding the overlapping nature of semantic and conceptual systems (e.g., de Groot, 2000; Pavlenko, 2000).

retrieval. It is assumed that in most language tasks the semantic system spreads its activation to both lexicons regardless of the target language. Thus, the flow of activation from the semantic system is target-language nonspecific regardless of the language in which a task is being performed (Costa & Caramazza, 1999; de Bot & Schreuder, 1993; Green, 1986, 1998; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Poulisse & Bongaerts, 1994). For instance, participants highly proficient in both Spanish and Catalan showed longer latencies for naming a target picture such as a "table" when the distractor was semantically related to the target (e.g., *chair* in Catalan [*cadira*] or Spanish [*silla*]) than when the distractor was semantically unrelated to the target. Similar findings have also been observed in Dutch–English bilingual speakers (Hermans et al., 1998). Costa, Miozzo, and Caramazza (1999) explained the crosslinguistic semantic interference by arguing that the semantically related word activates the corresponding concept, which spreads its activation to the lexical items in both languages. Double activation of the concept of an item can also result in faster naming (i.e., facilitation) when the picture of the item is presented with the written translation of the item (e.g., *taula* [*table* in Catalan]; Costa et al., 1999).

From the previous discussion, it can be assumed that a shared semantic system is theoretically connected directly to both lexicons (de Groot, 1992; Kroll & Stewart, 1994) and spreads activation to both lexicons (Costa & Caramazza, 1999; de Bot & Schreuder, 1993; Green, 1986, 1998; Hermans et al., 1998; Poulisse & Bongaerts, 1994). Given that semantic-based naming treatments in monolingual aphasia have resulted in successful improvements in trained and untrained items, the present study examined the effect of semantic naming treatment on crosslinguistic generalization of trained and untrained items in two languages. The extent to which premorbid proficiencies affect those generalization patterns is also addressed. Our specific predictions were as follows:

1. A semantic-based treatment focused on strengthening semantic representations will facilitate access to phonological representations for the trained items in the trained language (e.g., English: *apple*).
2. Generalization to the semantically related items in the trained language (e.g., English: *orange*) will occur. As previously discussed, treatments based on models of lexical processing that emphasize the underlying basis of lexical processing have been successful in facilitating generalization to untrained semantically related items (Drew & Thompson, 1999; Kiran & Thompson, 2003). Therefore, training lexical access by strengthening semantic features should potentially increase the level of activation of semantically related neighbors, thereby facilitating their retrieval during presentation of the corresponding picture stimuli.
3. Generalization to the translation of the trained item in the untrained language (e.g., Spanish: *manzana*) will occur because phonological representations of targets in both languages access a common semantic representation (de Groot, 1992), and semantic activation is thought to activate the phonological representations of both lexicons (e.g., Costa et al., 1999). Hence, the lexical forms in the target language will also activate translation equivalents in the nontarget language. Consequently, repeated exposure to targets as a function of treatment will facilitate phonological access to untrained translation equivalents in the nontarget language.
4. Generalization to the semantically related target in the untrained language (e.g., Spanish: *naranja*) will occur as a natural consequence of Predictions 2 and 3. Specifically, semantically related targets of the trained words become active when the target is active. Also, because phonological activation is hypothesized to be target-language nonspecific, trained and semantically related untrained words in untrained languages will also receive activation through the course of treatment.
5. No changes in a semantically unrelated control set (e.g., English: *boat;* Spanish: *vaca* [*cow*]) will occur, because these items should not be influenced by semantically based treatment.

Method

Participants

Three participants (P1, P2, and P3) with bilingual aphasia were recruited from local area hospitals. Several participant selection criteria were met in order for these individuals to be involved in the experiment: (a) diagnosis by a neurologist of a stroke in the left hemisphere (encompassing the gray/white matter in and around the perisylvian area) confirmed by a CT/MRI scan; (b) onset of stroke at least 9 months prior to participation in the study; (c) right-handed prior to stroke; (d) bilingual speakers of English and Spanish who reported being "functional" in both languages in most situations prior to their stroke; (e) relatively equal performance in both languages following their stroke; (f) adequate hearing, vision, and comprehension to engage fully in testing and treatment; and (g) stable health status. Age (range = 53–56 years) and years of education (P1 = 10 years; P2 and P3 = 12) were similar across participants. Please note that P2 was 8 months postonset at time of enrollment, but his baselines were deemed stable enough to begin treatment. See Table 1 for demographic details for the 3 participants.

Table 1. Demographic data, language history, and language proficiency ratings across languages for all participants.

Demographic information						Language history and proficiency				
Participant	M/F	Age	Education	Etiology	MPO	Family/Social	Work	Reading/Writing	Self-ratings (E/S) (1–7)	BPR
1	F	53	10 years (Mexico)	Left MCA CVA	9	Spanish only until 21 years Prior to CVA 100% English at home with husband Spanish and English with grown children Spanish only with brother English and Spanish with friends	**Factory:** 50% English 50% Spanish	Educated in Spanish Continued to write in Spanish (letters, lists) Learned and used English Read English and Spanish materials	Speech: 6/7 Comp: 7/7 Reading: 7/7 Writing: 7/7	1.08
2	M	53	12 years (U.S.)	Left MCA CVA	8	Both languages from birth Prior to CVA, mostly English with mother (bilingual) 100% English at home with wife No Spanish with friends	**Surveyor:** 70% English 30% Spanish	Educated in English No Spanish training Read and wrote primarily in English at home and at work	Speech: 7/5 Comp: 7/6 Reading: NA Writing: NA	0.79
3	F	56	12 years (U.S.)	Left MCA CVA	9	Both languages from birth Prior to CVA, 80% English and 20% Spanish (with husband) at home Spanish only with mother-in-law No Spanish with friends	**Retail:** 70% English 30% Spanish	Educated in English No Spanish training Read and wrote primarily in English at home and at work	Speech: 7/3 Comp: 7/5 Reading: NA Writing: NA	0.57

Note. M = male; F = female; MPO = months postonset; E = English; S = Spanish; Comp = Comprehension; MCA = middle cerebral artery; CVA = cerebral vascular accident; BPR = bilingual proficiency ratio.

Language proficiency levels. A number of methods were used to characterize each participant's language history and to estimate premorbid language-use patterns immediately prior to stroke. Each participant was interviewed and asked to complete a language-use questionnaire (Muñoz, Marquardt, & Copeland, 1999). Because self-reports have limitations (Hamers & Blanc, 2000; Romaine, 1995), at least one family member familiar with the participant's language acquisition and use was interviewed to corroborate information provided by the participants. Questions focused on the manner and time of acquisition for both languages as well as use patterns over time, with an emphasis on use and proficiency immediately prior to the cerebrovascular accident. Of interest was what languages were used at home, in social situations, and at work, and in what modalities. Furthermore, participants and family members rated premorbid proficiency in speech and comprehension in informal situations as well as reading and writing on a 7-point scale, with responses ranging from 1 (*not fluent*) to 7 (*native proficiency*). Proficiency in informal situations was used as the primary measure because the stimuli in the current study are common, concrete items. Language history, use patterns, and proficiency ratings were used to estimate premorbid proficiency in both languages.

As indicated in Table 1, P1 moved to United States from Monterrey, Mexico, when she was 21 years old; she acquired English as an adult. P1 appeared to be relatively balanced across languages in terms of use and proficiency, as she used English 100% of the time at home with her monolingual husband but used Spanish with one of her grown children, with her brother, and with friends. At work she used Spanish and English equally, and she read and wrote in both languages, even though she was only formally educated in Spanish. P2's and P3's families were from Mexico, and they reported that their Spanish was influenced by Mexican Spanish and Spanish spoken in central Texas. P2 and P3 exhibited more use and proficiency in English in all contexts and modalities, as they were both educated in English. They did not learn to read or write in Spanish, and primarily used English at home, at work, and in social situations.

To characterize premorbid language proficiencies further, the ratings for speech and comprehension provided in the interviews were used to calculate a bilingual proficiency ratio (BPR; BPR = Spanish comprehension + Spanish production/English comprehension + English production). The BPR for each participant was compared against those reported previously in normal Spanish–English bilinguals who fell into one of three proficiency groups: English dominant, Spanish dominant, or relatively balanced (Edmonds & Kiran, 2004). For example, P1's BPR (1.08; e.g., 7 [Spanish comprehension] + 7 [Spanish production]/7 [English comprehension] + 6] English production] = 1.08) most closely resembled the balanced bilingual group BPR (0.99), whereas P2's (0.79) and P3's (0.57) BPRs most closely resembled the English dominant group (0.88; see Table 1). The BPR provides additional estimates of proficiency level, corroborates reported language-use patterns, and allows for future comparisons across participants with different language proficiencies.

Participants' language abilities were examined using four tests. The Western Aphasia Battery (WAB; Kertesz, 1982) assessed aphasic symptoms and severity in English. Subtests from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992) were administered in English to determine semantic processing abilities pertinent to naming, the target skill investigated in this study. The Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 2001) examined naming abilities in both languages. The Bilingual Aphasia Test (BAT; Paradis, 1987) examined poststroke proficiency levels in each language. The semantic subtests of the BAT were administered in both languages, but the number of items (n = 5) in each subtest was small so that testing of the semantic system in both languages was limited for pre- and posttesting.

Performance on language tests in English. Aphasia quotient (AQ) results of the WAB showed that P1 and P3 presented with moderate aphasia (P1 AQ = 67.5; P3 AQ = 61.3) characterized by nonfluent speech, impaired comprehension, and naming deficits, with relatively spared reading comprehension of single words and phrases, whereas P2 presented with severe aphasia (P2 AQ = 27.0). Additionally, P2 exhibited characteristics consistent with apraxia of speech, including effortful speech with groping articulation and variable articulation errors.

On the PALPA, P1 and P2 performed above 90% on both spoken and written word to picture matching, whereas P2 was impaired on spoken (52.5%) and written word (77.5%) to picture matching. However, all participants demonstrated impairments in judging auditory and written word synonyms, with scores ranging from 48.3% to 73.3%. These data indicated mild-to-moderate semantic impairments for P1 and P3, with more severe impairments for P2 (see Table 2).

Performance on language tests in English and Spanish. All participants showed equal levels of naming performance across languages as measured by the BNT. Participant 1 was moderately impaired (P2 English = 41.7%, Spanish = 41.0%), while P2 and P3 showed more severe naming deficits (P1 English = 1.7%, Spanish = 0%; P2 English = 23.3%, Spanish = 18.3%). See Table 3 for results.

Table 2. Pre- and postlanguage performance on tests administered in English only (WAB; Kertesz, 1982, and PALPA; Kay et al., 1992).

Test	Participant 1		Participant 2		Participant 3	
	Pre	Post	Pre	Post	Pre	Post
Western Aphasia Battery (WAB)						
Spontaneous Speech (%)	60.0	65.0	**20.0**	**40.0**	65.0	70.0
Auditory Comprehension (%)	79.5	88.5	**47.0**	**61.5**	**74.5**	**87.5**
Repetition (%)	65.5	74.0	**27.0**	**38.0**	34.0	44.0
Naming (%)	**70.0**	**81.0**	**25.0**	**53.0**	68.0	73.0
Aphasia quotient (%)	67.5	74.7	**27.0**	**38.0**	61.3	68.9
Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)						
Spoken Word–Picture Matching (%)	92.5	97.5	**52.5**	**92.5**	95.0	95.0
Written Word–Picture Matching (%)	92.5	97.5	**77.5**	**95.0**	93.0	95.0
Auditory Synonym Judgments (%)	73.3	81.7	48.3	DNT	68.0	72.0
Written Synonym Judgments (%)	70.0	70.0	66.7	76.7	73.0	73.0

Note. Changes exceeding 10% are highlighted in bold. DNT = did not test.

Table 3. Pre- and posttreatment performance on tests administered in English and Spanish for all participants (BNT; Kaplan et al., 2001) and BAT (Paradis, 1987).

	Participant 1				Participant 2				Participant 3			
	English		Spanish		English		Spanish		English		Spanish	
Test	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Boston Naming Test (BNT)	41.7	48.3	**41.0**	**55.0**	**1.7**	**35.0**	0.0	1.67	23.3	33.3	**18.3**	**30.0**
Bilingual Aphasia Test (BAT)												
Pointing (%)	100	100	100	100	**50.0**	**100**	**30.0**	**60.0**	100	100	100	100
Semicomplex Commands (%)	**80.0**	**100**	**60.0**	**100**	**8.3**	**25.0**	**0.0**	**25.0**	**50.0**	**100**	90.0	90.0
Verbal Auditory Discrimination (%)	83.3	83.3	**66.7**	**100**	**61.1**	**72.2**	50.0	50.0	83.0	89.0	94.0	89.0
Judgment of Words/Nonwords (%)	**56.7**	**90.0**	**66.7**	**100**	**20.0**	**60.0**	0.0	0.0	90.0	97.0	97.0	93.0
Naming (%)	60.0	66.7	93.3	100	**0.0**	**42.9**	0.0	DNT	88.0	95.0	74.0	79.0
Word Repetition (%)	93.3	96.7	96.7	96.7	76.7	73.3	66.7	DNT	**57.0**	**77.0**	**57.0**	**77.0**
Semantic Categories (%)	**100**	**80.0**	**100**	**80.0**	**60.0**	**80.0**	**100**	**60.0**	**100**	**60.0**	**80.0**	**60.0**
Semantic Opposites (%)	**20.0**	**40.0**	**10.0**	**50.0**	**0.0**	**70.0**	10.0	10.0	30.0	20.0	**30.0**	**10.0**
Semantic Acceptability (%)	100	100	90.0	100	70.0	60.0	50.0	40.0	90.0	100	**80.0**	**100**
Synonyms (%)	80.0	80.0	**60.0**	**80.0**	**20.0**	**100**	**20.0**	**0.0**	**20.0**	**0.0**	**40.0**	**60.0**
Antonyms I (%)	80.0	80.0	**100**	**60.0**	**20.0**	**60.0**	**0.0**	**40.0**	**40.0**	**60.0**	40.0	40.0
Antonyms II (%)	**60.0**	**100**	80.0	80.0	**20.0**	**40.0**	**60.0**	**0.0**	**40.0**	**0.0**	**80.0**	**40.0**
Reading Words (%)	80.0	70.0	90.0	100	0.0	0.0	0.0	0.0	90.0	90.0	0.0	10.0
Reading Sentences (%)	**20.0**	**50.0**	**40.0**	**70.0**	**0.0**	**16.7**	0.0	DNT	0.0	0.0	0.0	0.0
BAT—Part C												
Recognition of words (Spanish to English) (%)	100	100	NA	NA	**20.0**	**100**	NA	NA	100	100	NA	NA
Recognition of words (English to Spanish) (%)	100	100	NA	NA	0.0	10.0	NA	NA	**100**	**80**	NA	NA
Translation of words (Spanish to English) (%)	50.0	60.0	NA	NA	0.0	0.0	NA	NA	**10.0**	**40.0**	NA	NA
Translation of words (English to Spanish) (%)	60.0	60.0	NA	NA	0.0	0.0	NA	NA	**20.0**	**50.0**	NA	NA

Note. Changes exceeding 10% are highlighted in bold. DNT = did not test; NA = not applicable.

Most subtests of interest of the BAT were administered, including Part C, which evaluates recognition and translation of words across languages, because the current study investigated crosslinguistic generalization. Results of the BAT revealed that performance levels in English and Spanish were generally equivalent within each participant across languages, an important criterion for inclusion in the study. For P1, accuracy was equal to or less than 50% for Semantic Opposites and Reading Sentences in both English and Spanish. On all other BAT subtests, performance was >50%, indicating milder deficits. On the translation subsection of the BAT, P1 was able to recognize all of the words across languages from Spanish to English, and vice versa, and she could translate 50% of the words from Spanish to English and 60% from English to Spanish. P2 was more severely impaired in both languages, with accuracy equal to or less than 50% on 10 of 14 subtests in English and 11 of 14 subtests in Spanish. For P3, accuracy was equal to or less than 50% on Semantic Opposites, Synonyms, and Antonyms in English and Spanish and on Semicomplex Commands in English only. Similar to P1, P3 was able to recognize all words across languages in both directions, but was more impaired in translating in both directions (see Table 3 for results).

In summary, all participants demonstrated relatively equal performance on pretreatment testing across languages. Overall, P2 (premorbidly more proficient in English) was more severely impaired than P1 (premorbidly equally proficient) and P3 (premorbidly more proficient in English) on most tests administered.

Stimuli

Development of treatment stimuli. From an original corpus of 200 words that varied across semantic categories, 150 were selected based on the following criteria. Cognates (e.g., *elephant* and *elefante*) and words with at least 50% phonetic similarity (e.g., *cat* and *gato*) were eliminated from the set. Words between one and four syllables (English average = 1.53; Spanish average = 2.58) were then chosen. For P2, high- to moderate-frequency words were selected in each language (Frances & Kučera, 1982; Juilland & Chang-Rodriguez, 1964) such that the average frequency for both languages (English = 53.86, *SD* = 107.35; Spanish = 58.50, *SD* = 126.47) was matched as determined by a paired *t* test, $t(19) = 1.66$, $p = .114$. Confrontation naming of these items in both languages was performed on a group of 23 normal bilingual individuals; these results have been reported elsewhere (Edmonds & Kiran, 2004). Additional stimuli were developed for P1 and P3, as they were able to name many of the items from the original set. The stimuli were selected using the same procedures as with P2. Moderate- to low-frequency words were selected in each language (Frances & Kučera, 1982; Juilland & Chang-Rodriguez, 1964) such that the average frequency for both languages was matched as determined by a paired *t* test (P1: English = 8.95, Spanish = 3.3; $t[19] = 1.96$, $p = .065$; P3: English = 16.7, Spanish = 22.3; $t[19] = 1.53$, $p = .142$). The picture stimuli were chosen from Art Explosion Software (NOVA, Inc.) and modified to equal approximately 4 × 6 in. and were centered on 8.5 × 11 in. white paper.

For each participant, six stimulus sets were created: English Set 1 (e.g., *apple*), Spanish Set 1 (e.g., *manzana*), English Set 2 (e.g., *orange*), Spanish Set 2 (e.g., *naranja*), English control set (e.g., *boat*), and Spanish control set (e.g., *vaca* [*cow*]). All sets except the control sets ($n = 5$ each) contained 10 items each, resulting in 50 different items for each participant. Set 2 was semantically related to Set 1 in each language. Ratings of the semantic relatedness of pairs of words were completed by normal bilinguals as described previously (Edmonds & Kiran, 2004). Word pairs were all category coordinates (e.g., *apple* and *orange*) except for one superordinate–category member pair (*fish–shark*). As much as possible, control sets contained items that were semantically unrelated to English and Spanish Sets 1 and 2 and to each other; however, there were some exceptions for P3, which will be discussed in the Results section. See the Appendix for a list of stimuli for all participants.

Development of semantic features for treatment. For each item (e.g., *apple*), five semantic features referring to the superordinate category (e.g., *fruit*), function (e.g., *provides nutrition*), general characteristic (e.g., *is sweet*), physical characteristic (e.g., *has skin/peel*), and location (e.g., *found in refrigerator/produce section*) were developed prior to treatment for each item. To increase the functional value of the treatment, each participant generated his or her own personal association with each item as the sixth feature during the first few weeks of treatment with assistance from the clinician as needed. In addition to the six target semantic features, six distractor features for each item were created.

Design

A single-subject experimental multiple baseline design across participants and behaviors (Connell & Thompson, 1986; McReynolds & Kearns, 1983) was used to examine generalization of semantically related items within each language and across languages. As required for a multiple baseline design across behaviors, the purpose of the treatment was to demonstrate replication of a single treatment (i.e., semantic treatment) as being effective in improving naming skills across both languages in 1 participant. In addition to varying the number of baselines, the order of stimuli sets and language was counterbalanced across patients. Therefore, P1 and P3 were initially trained on Spanish Set 1 (e.g., *apio*), with generalization monitored for Spanish Set 2 (e.g., *repollo*), English Set 1 (e.g., *celery*), and English Set 2 (e.g., *cabbage*). P2 was first trained on English Set 1 (e.g., *apple*), and generalization to English Set 2 (e.g., *orange*), Spanish Set 1 (e.g, *manzana*), and Spanish Set 2 (e.g., *naranja*) was assessed. Only if no generalization to crosslinguistic translations was noted when the accuracy of trained items reached criterion (80%) would treatment be shifted to the untrained language. Such a design allowed examination of generalization to semantically related items within the same language in addition to examining crosslinguistic generalization. Participants received different stimuli sets that were selected in accordance with their level of naming deficits. No generalization was expected on the control set of stimuli in either language.

Baseline Measures

During baseline sessions, naming of the 50 examples was tested through a confrontation naming task. Participants were shown each picture and were instructed to name each item. Stimuli were presented in language blocks with the order of stimuli pseudorandomized within each block to ensure that items from the same category (e.g., *apple* and *orange*) were not presented sequentially. Prior to presentation of stimuli in each language, the bilingual clinician conversed with the participant for a minimum of 5 min to ensure that participants were aware of the target language. The order of presentation of languages was counterbalanced across sessions.

Oral responses were considered correct if they were clear and intelligible productions of the target item. Self-corrected responses, dialectical differences, distortion/substitution, or addition/omission of one vowel or consonant were allowed. All other responses, including (a) cross-linguistic correct responses (i.e., correct name for item but in the wrong language [e.g., *drill/taladro*]), (b) cross-linguistic semantic responses (e.g., *naranja* [*orange*]/*fruit*),

(c) crosslinguistic unrelated words (e.g., *anillo* [*ring*])/*rake*), (d) semantic errors (*leg/arm*), (e) unrelated word responses (e.g., *garlic/radio*), (f) neologisms and perseverations (syllable, word, or neologism produced three or more times during the specific probe session), and (g) no responses (NR)/"I don't know" (IDK) responses were all scored as incorrect. Neologisms and perseverations were combined into a single category because P2 exhibited many instances of perseveration on neologisms. A category for phonemic errors was not created because no participant produced phonemic errors.

Baselines were considered stable as long as no more than 30% variability and no more than 10% rise in the last baseline were seen. These criteria have been used in other published studies (Kiran, 2005; Kiran & Thompson, 2003) and accommodate for fluctuations in performance, given that patients actually saw each picture twice (in English and Spanish) within a given baseline session.

Treatment

To facilitate access to naming of the trained items, semantically based treatment methodologies developed by Boyle and Coelho (1995) and Kiran and Thompson (2003) were employed. Treatment was provided in one language two times per week for 2-hr sessions. The first hour of the second session was dedicated to probes, and treatment was provided the second hour. Participants performed five treatment steps that emphasized semantic feature attributes of that particular example. First, participants were required to name the example. Irrespective of whether the target was named accurately or not, the clinician said the name of the object and then set a card with the written form below the picture. Participants were then provided with a set (n = 12) of written target semantic feature cards (n = 6) and distractors (n = 6), and were instructed to select the target semantic features for each example. For each correct semantic feature, the clinician reinforced whether the selection belonged to the six attribute types (e.g., superordinate label, function, characteristic). If participants did not understand the instructions or terminology, the clinician provided additional information or modeled what was expected. Over time, participants were encouraged to respond more independently. Following the selection of the related features, the picture was turned over and the participants were asked 12 yes–no questions regarding the features (e.g., "Is it a fruit?" or "Is it found on the roof?") and were required to accept or reject the features as being applicable to the target example. Finally, the picture was presented again, and the participants were required to name it. An attempt was made to practice all items at least one time per week. All participants were shown all 10 trained pictures at the end of each session and were asked to name the pictures. During treatment, the bilingual clinician always remained in the target language regardless of which language was used by participants.

Treatment Probes

Throughout treatment, naming probes (n = 50) like those presented in the baseline condition were administered at the beginning of every second treatment session to assess naming of the trained and untrained items. P1 exhibited nervousness during testing situations, so two modifications were made to the protocol. First, naming probes for the unrelated control items were administered every third session to decrease the number of items tested. Second, online response transcription was conducted by a reliability scorer (the first author) behind a one-way mirror. As in baselines, the order of presentation of items was pseudorandomized during each probe presentation, and the order of language presented was counterbalanced across sessions.

Data Analysis

Responses to naming probes, coded in the same way as in baseline, served as the primary dependent measure in the study. Treatment was discontinued when naming accuracy of 8/10 items (80% accuracy) was observed for two consecutive sessions or when a total of 20 treatment sessions (10 probe sessions) were completed. For P2, Spanish treatment was discontinued when naming accuracy reached 100% because the participant had been in treatment for 66 sessions. Generalized naming to the untrained examples was considered to have occurred when levels of performance increased by at least 40% over baseline levels. The extent to which changes from baseline to treatment phases were statistically reliable was determined through a time series analysis using the C statistic (Tryon, 1982). Maintenance testing of the 50 items was conducted 1 and 2 months posttreatment for P1 and P2. P3 received maintenance testing 1 month posttreatment. Procedures and analysis were identical to those used during baseline testing.

Reliability

All of the baseline and probe sessions were recorded on videotape. Reliability for the independent variable (treatment) was conducted by a licensed bilingual speech-language pathologist who sat behind the two-way mirror and determined percentage of steps that were performed correctly by the clinician for 50% of all the treatment sessions. Point-to-point agreement indicated a reliability of 99%. Additionally, reliability was conducted on the dependent variable (naming responses) for 75% of the baseline and probe sessions. Point-to-point reliability agreement indicated 95% reliability. Any discrepancies went to the judgment of the clinician present in the room with the participant. Additionally, reliability was performed during error analysis by the second author on 20% of the naming errors with 100% agreement.

Results

Naming Accuracy

Results for all participants are presented in Figures 1–4 in a multiple baseline format showing the percentage of correct items for Spanish and English trained and semantically related untrained items and unrelated control items.

Participant 1. Following two baseline sessions, treatment was initiated on Spanish items (Spanish Set 1), which resulted in acquisition of those items to criterion (10%–90% accuracy; $C = 0.705$, $z = 2.926$, $p = .001$). Generalization was observed in the untrained semantically related items in Spanish Set 2 (10%–70% accuracy; $C = 0.705$, $z = 2.926$, $p = .001$). Notably, crosslinguistic generalization was observed in the untrained English translations in English Set 1 (10%–50% accuracy). Although generalization met the criterion, this improvement was variable as is evident from Figure 1 and was not statistically significant ($C = 0.363$, $z = 1.465$, $p = .071$). Finally, crosslinguistic generalization was also observed for untrained semantically related words in the untrained language (English Set 2) with accuracy improving from 10% to 60% ($C = 0.700$, $z = 2.905$, $p = .001$).

Generalization to English Set 1 was somewhat weak due to the variability in performance. Data analysis revealed that during probes, P1 oftentimes *initially* produced the correct lexical item in the nontarget language (e.g., for the target *gusano* in Spanish, she would first say *worm* but would eventually say "I don't know"). Thus, P1 could accurately retrieve a correct phonological form, but not always in the target language. For example, during the eighth probe session she incorrectly produced the translation equivalent of the target item in the wrong language 9/20 times. To characterize this phenomenon, correct lexical access irrespective of the target language was graphed (e.g., *worm*/*gusano* irrespective of target language) in addition to her existing data. As is evident from Figure 2, P1 achieved a high of 100% on the trained items and 90% on Spanish Set 2 and for both English sets. In other words, when trained only on Spanish Set 1, she was able to retrieve 90% of the other three sets of words when she was not penalized for crosslinguistic responses.

Performance on the unrelated control words in each language ($n = 5$) was measured every third probe session to reduce this participant's testing anxiety. As is illustrated in Figure 1, performance on unrelated words in Spanish improved from 0 to 40% accuracy (0/5–2/5 items correct) during treatment. For one item (*sobre* [*envelope*]), she typically provided a semantic error during baseline (*carta* [*letter*]), which was resolved during treatment. Another item (*bombero* [*fireman*]) was a word she was personally familiar with, and, hence, strived to retrieve the word during probes. The English control set, however, did not demonstrate any change as a function of treatment.

Maintenance probes conducted 1 month following treatment (Session 18) and 2 months following treatment (Session 19) revealed performance levels above baseline for all stimuli sets. Interestingly, performance levels on the untrained English stimuli were better maintained than the trained and untrained semantically related Spanish sets.

Participant 2. Following three baselines, P2 received naming treatment for English Set 1, which improved from 10% to 80% accuracy after 18 probe sessions ($C = 0.87$, $z = 4.20$, $p = .001$, see Figure 3). Additionally, naming of semantically related words (English Set 2) also improved to generalization criterion (10%–50% accuracy, $C = 0.846$, $z = 3.984$, $p = .001$). Performance on Spanish Set 1 and Set 2 did not change as a function of treatment, indicating that crosslinguistic generalization did not occur from English to Spanish. Treatment was then shifted to Spanish Set 2. Acquisition of trained items reached 100% accuracy (0%–100% accuracy, $C = 0.826$, $z = 4.898$, $p = .001$) after 15 probe sessions. Unlike the English treatment results, no within-language generalization to semantically related words was observed in Spanish (0%–20%; $C = -0.089$, $z = -0.530$, $p = .702$). Notably, with naming treatment for Spanish Set 2, performance on the untrained translations in English Set 2 improved from 30% accuracy at the end of the English Set 1 treatment phase to a high of 80% accuracy between Sessions 22 and 37 ($C = 0.877$, $z = 5.201$, $p = .001$). Crosslinguistic generalization was noted on the originally trained English Set 1, which, upon initiation of Spanish treatment, improved from 80% to 100% accuracy, a significant change from its treatment levels ($C = 0.880$, $z = 5.218$, $p = .001$). Unrelated control items ($n = 5$ items in each language) did not change as a function of treatment.

Maintenance probes conducted 1 month following treatment (Session 39) and at 4 months following treatment (Session 40) revealed performance above baseline levels for trained English Set 1 and generalized English Set 2. Performance on the five unrelated control items in English increased from 0% (0/5) to 40% (2/5) accuracy in the maintenance probes. Performance on trained Spanish Set 2 items was maintained at 50% accuracy 1 month following treatment, but fell to baseline levels 4 months after treatment. No changes were observed on untrained Spanish Set 1 items or Spanish unrelated control items.

Figure 1. Naming accuracy for Participant 1 on Spanish Set 1 (trained) and Spanish Set 2 (semantically related to Set 1), English Set 1 (translations of Spanish Set 1) and English Set 2 (semantically related translations). Control items in English and Spanish are illustrated in the same graph as English and Spanish Set 1 items.

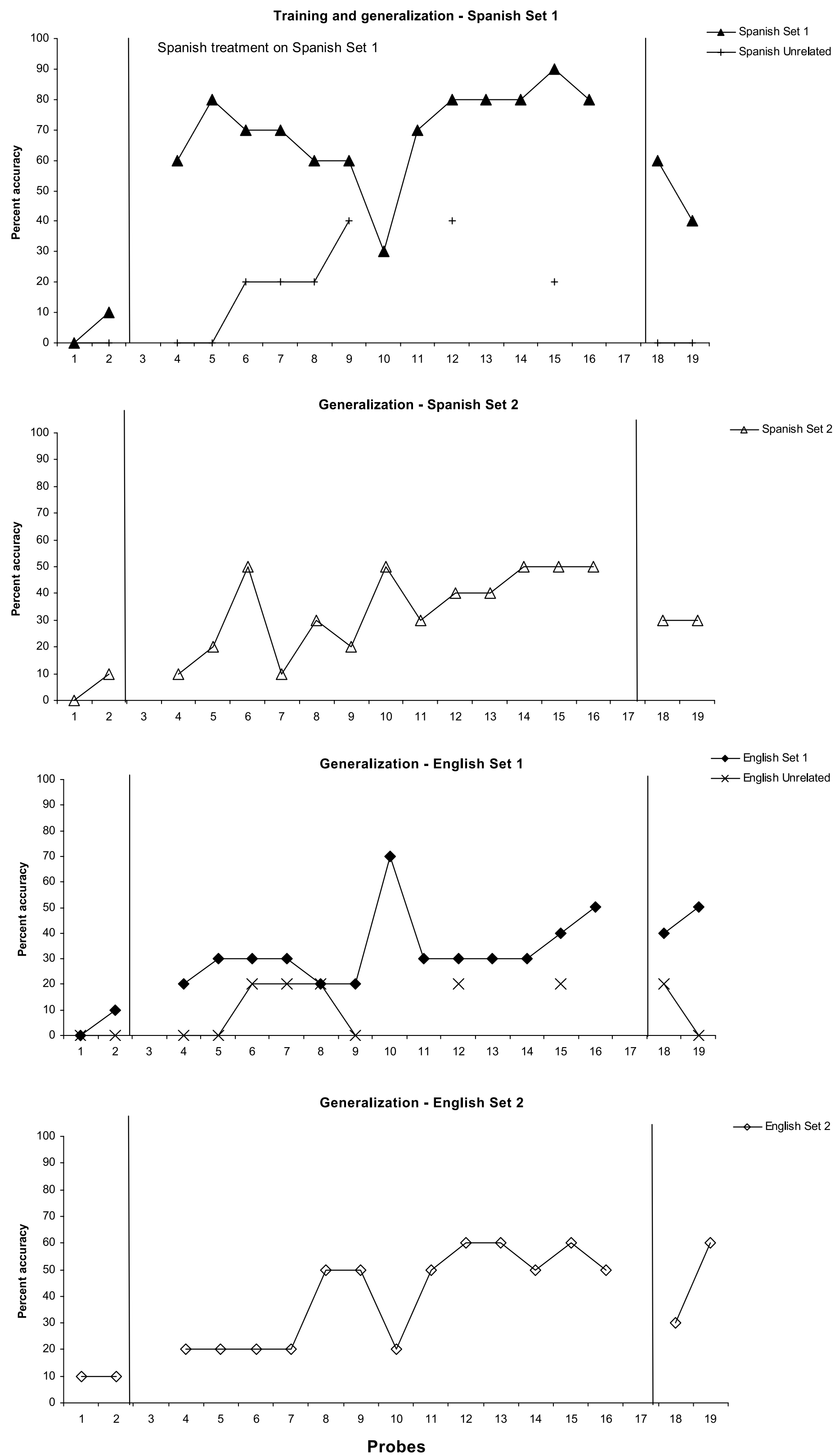


Figure 2. Naming accuracy for Participant 1 on Spanish Set 1 and Set 2 and English Set 1 and Set 2 when accurate lexical access irrespective of target language was calculated (e.g., *gusano* → worm = correct because correct response in nontarget language).

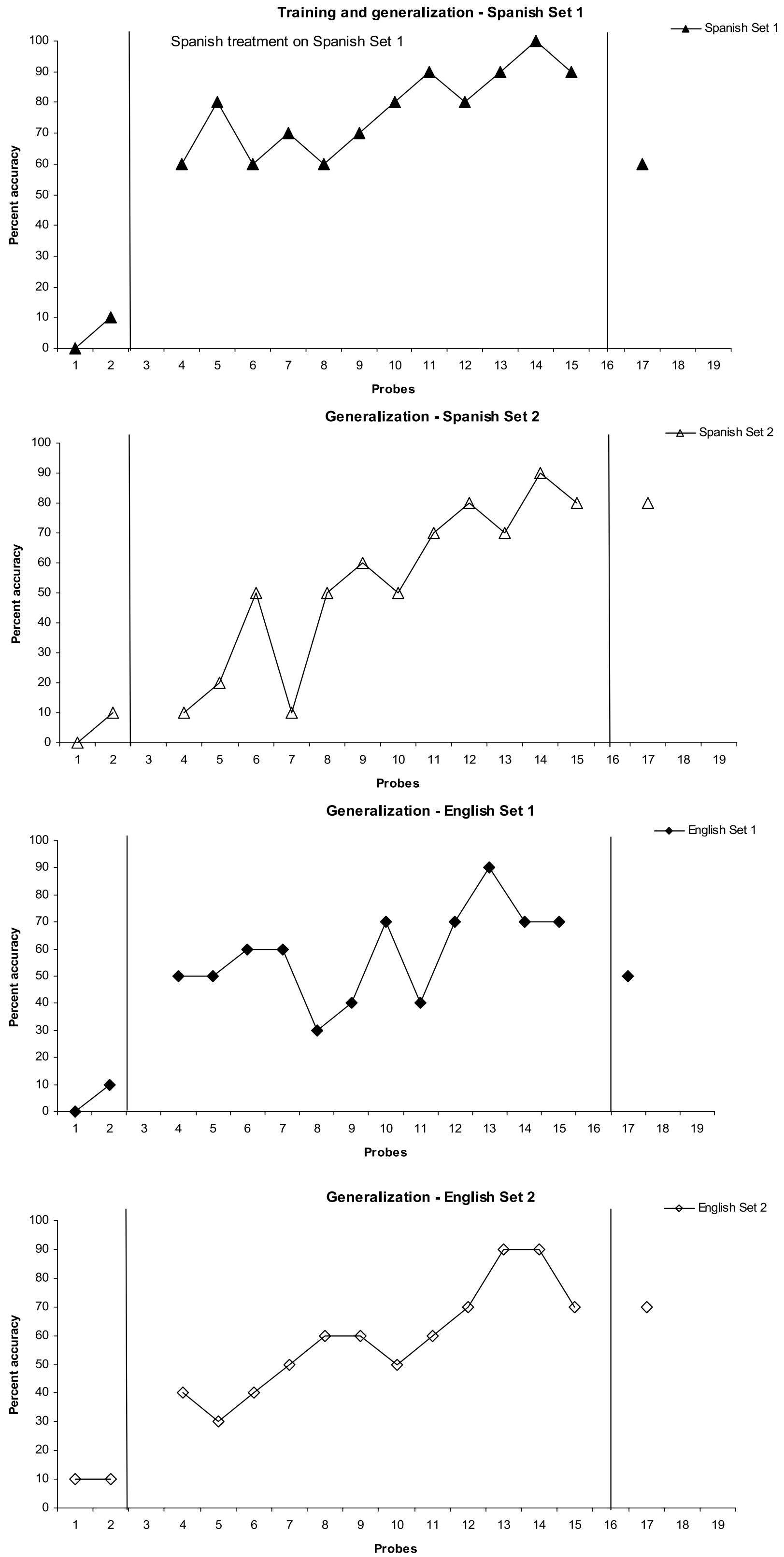
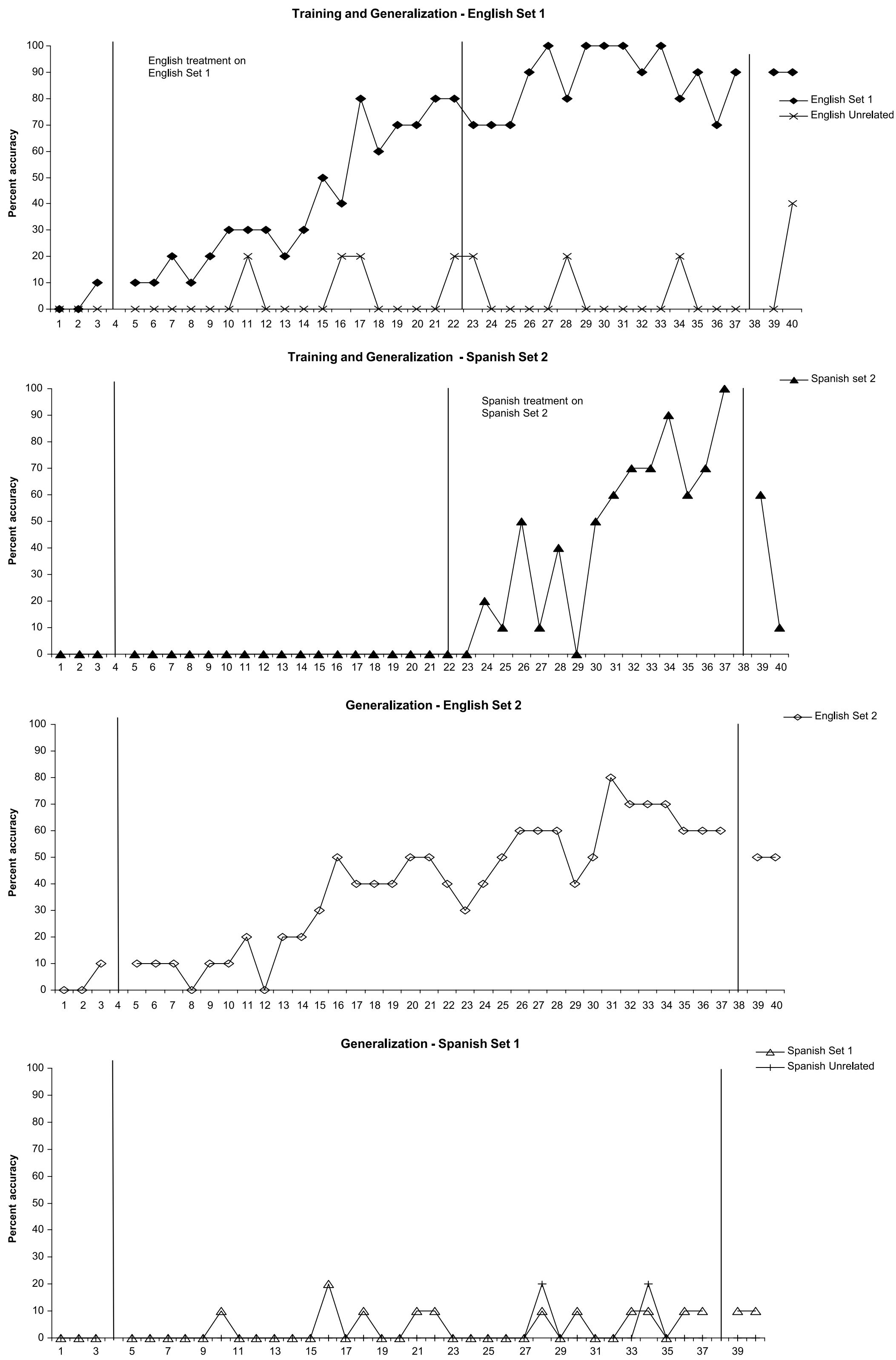


Figure 3. Naming accuracy for Participant 2 on English Set 1 (trained first) and English Set 2 (semantically related to Set 1) and Spanish Set 1 (translations of English Set 1) and Spanish Set 2 (trained second). Control items in English and Spanish are illustrated in the same graph as English and Spanish Set 1 items.



Participant 3. Following four baseline sessions, treatment was initiated on Spanish items (Spanish Set 1), which resulted in improvements of those items to criterion (20%–90% accuracy; $C = 0.590$, $z = 1.829$, $p = .033$; see Figure 4). However, no improvements were observed on the untrained semantically related items in Spanish Set 2 (30%–40% accuracy; $C = 0.000$, $z = 0.000$, $p = .500$). Crosslinguistic improvement was observed on untrained English translations (Set 1; 40%–70% accuracy). However, because the baseline exceeded our criteria, and generalization did not meet criterion ($C = 0.270$, $z = 0.839$, $p = .200$), these results are interpreted with caution. Finally, crosslinguistic generalization was observed for untrained semantically related words in the untrained language (English Set 2) with accuracy improving from 30% to 70%, a statistically significant improvement ($C = 0.562$, $z = 1.742$, $p = .040$).

Performance on the unrelated control words differed across languages. As is illustrated in Figure 4, performance on unrelated words in Spanish improved from 40% to 60% accuracy (2/5 to 3/5 items correct) during treatment, indicating no appreciable gains. However, the unrelated items in English improved from 40% (2/5) to 100% (5/5), an appreciable difference. As previously discussed, it was difficult to control for semantic category when selecting the control items for this participant because of her variable naming results. As a result, four of the five unrelated words formed semantically related pairs (e.g., necklace/*pulsera* [bracelet]), hence facilitating generalization. Maintenance probes conducted 1 month following treatment (Session 12) revealed levels consistent with final probe performances on Spanish Set 1 and generalized English Set 1 and English Set 2.

Error Analysis

Evolution of errors through the course of treatment was analyzed for all participants. For P1 and P3, all baseline responses were compared with an equal number of sessions at the end of treatment. Because P2 received treatment in both languages without formal baselines before Spanish treatment, errors produced during the first three and the last three sessions for each treatment condition (English and Spanish) were analyzed. English errors and Spanish errors were examined separately for all participants. The proportions of errors by type across participants are included in Table 4.

For P1, English baseline errors were primarily IDK/NR (63.0%) with a few unrelated (1.7%) and crosslinguistic (1.7%) errors. At the end of treatment, her IDK/NR responses reduced to 13.4% errors with few semantic (1.7%) and unrelated (1.7%) errors. In Spanish, her errors prior to treatment were IDK/NR (76.4%), which by the end of treatment reduced to 4.0% errors with some unrelated responses (4.0%). In both English and Spanish, there was a notable increase in crosslinguistic errors; that is, Spanish-for-English errors increased to 16.7% and English-for-Spanish errors increased to 15.6%.

For P2, baseline errors in English were predominantly perseverations/neologisms (36%) with unrelated responses (10.5%) as the only other error type. At the end of all treatment, errors were primarily semantic (7.0%) with no perseverations and few IDK/NR errors (2.5%). Spanish errors were also primarily perseveration/neologisms (21.6%) with some crosslinguistic errors (7.2%). There was little change in Spanish error types at the end of English treatment. However, by the end of Spanish treatment, perseveration/neologisms were drastically reduced (3.8%) with virtually no crosslinguistic errors (0.5%), while IDK/NR responses increased from 2.0% to 9.1%.

For P3, English baseline errors were primarily semantic (35.5%) and IDK/NR (20.4%) with a few neologistic (4.3%), unrelated (4.3%), and crosslinguistic (1.1%) errors. At the end of treatment, all error categories showed a reduction (semantic = 15.0%, IDK/NR = 12.9%, unrelated = 1.1%, neologisms = 0.0%) except for crosslinguistic errors (5.4%), which increased slightly. Spanish error patterns were similar to English error patterns. During baseline, errors were predominantly semantic (25.8%) and IDK/NR (23.7%) with some neologistic (7.2%), unrelated (4.0%), and crosslinguistic (6.1%) errors. At the end of treatment, all error categories showed a reduction (semantic errors = 6.2%, neologisms = 2.0%; IDK/NR = 20.0%; crosslinguistic = 5.0%).

Pre–Post Standardized Language Measures

All tests administered prior to initiation of treatment were reassessed upon completion of treatment (see Tables 2 and 3). Pre- and posttesting was not always conducted by the same clinician for each participant; however, all scoring was verified by the first author, who is bilingual. On the BAT (Paradis, 1987), only subtests that may be expected to improve with treatment are considered: Naming, Semantics, Comprehension (because participants answered questions during treatment), and Reading (because participants read feature cards). Only changes that exceeded 10% are discussed in the text, but all results are reported in Tables 2 and 3.

On tests administered in English only, P1 demonstrated a modest improvement on the WAB (67.5–74.7), with notable improvement on the Naming subtest. On the BNT, improvement was seen in Spanish but not in English. On the BAT, improvements were observed in both languages on Comprehension (semicomplex commands, judgment of real/nonwords) and Reading Sentences. Additional improvements were observed in Spanish on Verbal Auditory Discrimination and Synonyms. A decrease on posttest scores on the Semantic Categories subtest in both languages and on Antonyms I in Spanish occurred along

Figure 4. Naming accuracy for Participant 3 on Spanish Set 1 (trained) and Spanish Set 2 (semantically related to Set 1) and English Set 1 (translations of Spanish Set 1) and English Set 2 (semantically related translations). Control items in English and Spanish are illustrated in the same graph as English and Spanish Set 1 items.

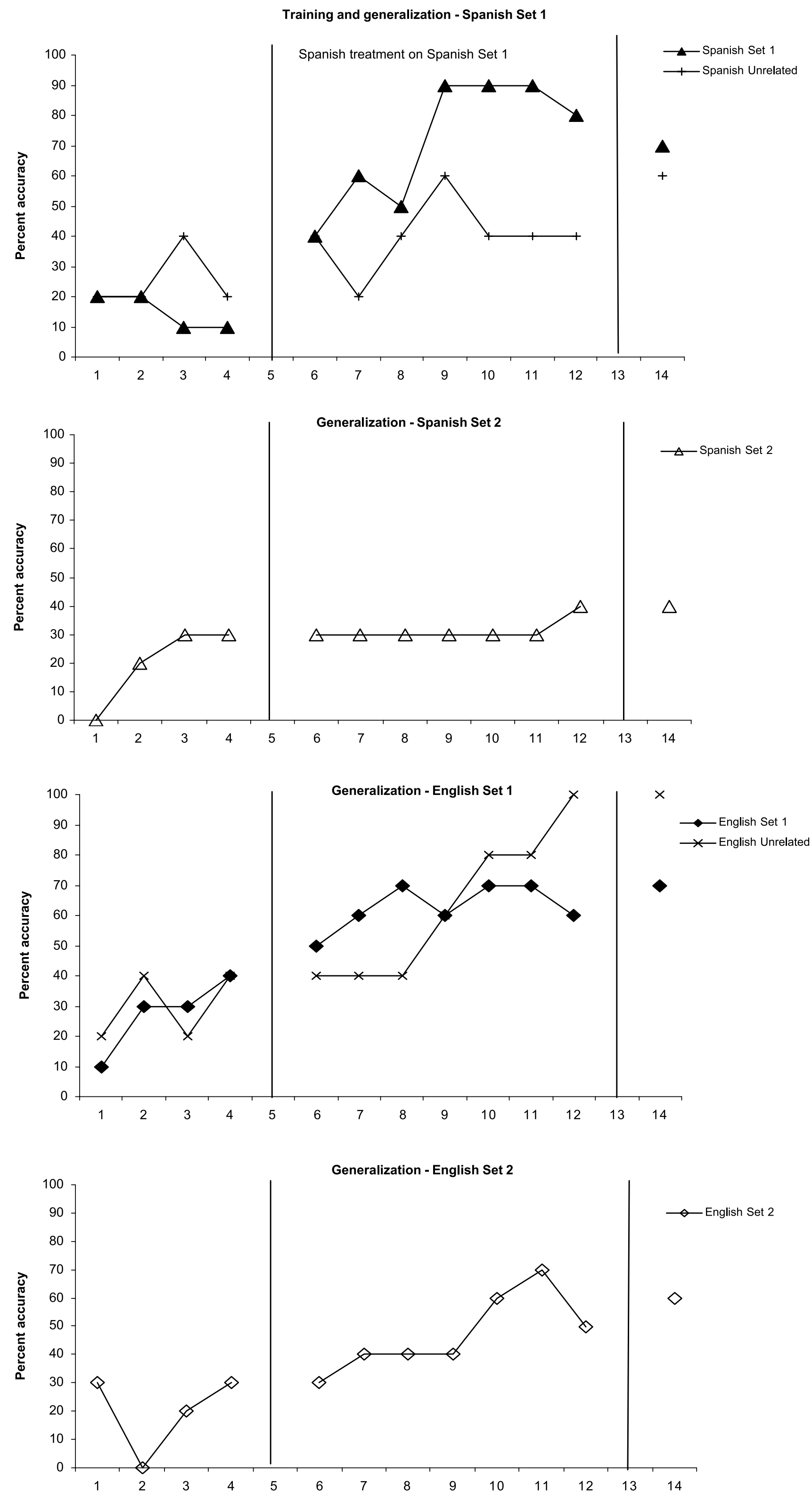


Table 4. Evolution of errors from pretreatment to posttreatment, reported in percentage of specific errors to total errors for all participants.

Error type	Participant 1		Participant 2				Participant 3	
	Spanish treatment		English treatment		Spanish treatment		Spanish treatment	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
English probes								
Perseveration/neologism	NA	NA	36	4.4	7.9	0	4.3	0.0
"I don't know"/no response	63.0	13.4	0	0.9	7.0	2.5	20.4	12.9
Unrelated	1.7	1.7	10.5	5.3	3.5	3.5	4.3	1.1
Semantic	0	1.7	0	5.3	5.3	7.0	35.5	15.0
Crosslinguistic	1.7	16.7	0	0	0	0.9	1.1	5.4
Total	66.4	33.6	46.5	15.9	23.7	13.9	65.6	34.4
Spanish probes								
Perseveration/neologism	NA	NA	21.6	24	19.7	3.8	7.2	2.0
"I don't know"/no response	76.4	4.0	0	0	2.0	9.1	23.7	20.0
Unrelated	0	4.0	0	0	0	2.0	4.0	0.0
Semantic	0	0	0	0	0	1.0	25.8	6.2
Crosslinguistic	0	15.6	7.2	3.8	5.3	0.5	6.1	5.0
Total	76.4	23.4	28.8	27.8	27.0	16.4	66.8	33.2

with an increase in Semantic Opposites and Antonym II subtest scores in both languages. This variable pattern of change in Semantic Opposites subtests across languages on the BAT was seen in all participants and can be likely attributed to the small number (n = 5) of items in each subtest.

On tests administered in English only, P2 demonstrated overall improvement on the WAB AQ (27–38) with improvements in Spontaneous Speech, Auditory Comprehension, Repetition, and Naming. On the PALPA, notable improvements were observed in Spoken Word–Picture Matching and Written Word–Picture Matching. On the BNT, improvement was seen in English but not in Spanish. On the BAT, improvements were observed in both languages on two Auditory Comprehension subtests and on one Semantics subtest. Improvements in English only were observed on Naming, one Reading, two Auditory Comprehension, and two Semantics subtests. No subtest improved in Spanish only, but two Semantic subtest scores decreased. P2 improved on Recognition of Words from Spanish to English on Part C of the BAT.

On tests administered in English only, P3 demonstrated a modest improvement on the WAB (61.3–68.9) with notable improvements on the Auditory Comprehension subtest. Improvements were observed in both languages on the BNT. On the BAT, comprehension of semicomplex commands and antonyms improved in English, and semantic acceptability and synonyms improved in Spanish. As seen in P1 and P2, some decrease in Semantic subtests was also observed in both languages. Part C revealed a notable improvement of P3's ability to translate words from English to Spanish, and vice versa.

Discussion

The aim of the present experiment was to examine crosslinguistic generalization patterns systematically in three English–Spanish bilingual individuals with aphasia. Results revealed crosslinguistic generalization in all participants. However, the patterns of generalization differed across participants, and these differences appear to be related to premorbid language proficiency levels.

P1 was premorbidly equally proficient in English and Spanish. Training in Spanish resulted in recovery of trained items and within-language generalization to semantically related items. Additionally, crosslinguistic generalization was observed in untrained semantically related items with marginal improvement to the untrained English translation equivalents of trained Spanish items. As all sets of items demonstrated improvements, treatment was not provided in English, and it remains to be seen if English-to-Spanish generalization would have occurred. As predicted, no appreciable gains were observed in the Spanish and English control sets, indicating that observed improvements were due to treatment.

P2 and P3 were both premorbidly more proficient in English than in Spanish. Whereas P2 first received treatment in English followed by Spanish, P3 received treatment in Spanish only. With English treatment, P2 showed within-language generalization to semantically related items within the more proficient language but no crosslinguistic generalization. However, when treatment was shifted to Spanish (the less proficient language), improvement of trained items was observed, but within-language generalization to semantically related untrained items was not observed. Importantly, with Spanish treatment, crosslinguistic generalization was observed for English translation equivalents and corresponding semantically related items. These results were replicated in P3, who was also more proficient in English. Specifically, treatment in Spanish did not result in within-language generalization but facilitated crosslinguistic generalization to English Set 2 with marginal improvements in English Set 1. The main finding of this study was that crosslinguistic generalization was observed in all participants. P1, who was premorbidly equally proficient across languages, was trained in Spanish and showed generalization to English. P2 and P3 were premorbidly more proficient in English, but when trained in Spanish showed generalization to English.

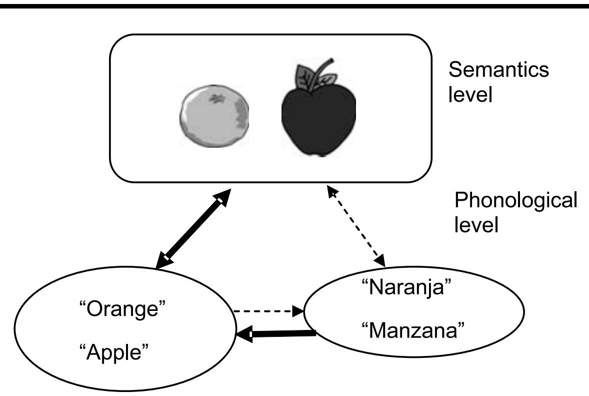
Several findings are consistent with our initial predictions. First, treatment focused on strengthening semantic representations resulted in improvements of trained items in the trained language for all participants. Furthermore, when treatment was provided in the first or premorbidly dominant language, generalization to semantically related items in the trained language was observed. These observations supplement the increasing body of evidence reporting the beneficial effects of a theoretically motivated semantic treatment to facilitate lexical retrieval and generalization in individuals with aphasia (Drew & Thompson, 1999; Kiran & Thompson, 2003). As a corollary, training the less proficient language in 2 participants did not facilitate generalization to semantically related items within the trained language, an interesting finding that merits further discussion.

The generalization patterns for all participants can be explained by extending the theoretical assumptions of bilingual lexical access models to accommodate for language recovery in bilingual aphasia. One model that accommodates all 3 patients' results is de Groot's mixed model (de Groot, 1992; de Groot et al., 1994), which allows flexibility in the strengths of connections between the semantic system and each lexicon as well as between lexicons themselves based on relative proficiencies across languages. This model proposes equally strong connections from conceptual memory to both lexicons and between both lexicons in an equally proficient bilingual individual such that treatment in one language should result in generalization within the trained language as well as to the untrained language. P1's results, in which extensive improvements within and across languages were observed with only one set of items treated in one language, are consistent with the mixed model's proposed connection strengths for equally proficient bilingual individuals. This finding is consistent with previous studies that found that activation of a target in one language also activates semantically related items in the target language as well as in the translation equivalents of the target language and semantically related items in another language (e.g., Costa et al., 1999; Hermans et al., 1998). The current findings are also consistent with the well-accepted assumption that the conceptual system spreads its activation to both lexicons in most language tasks regardless of the target language (e.g., Costa & Caramazza, 1999; de Bot, 1992; Green, 1986, 1998; Hermans et al., 1998; Poulisse & Bongaerts, 1994).

P2 showed the expected within-language generalization patterns when treated in English, his dominant language, but no crosslinguistic generalization to his premorbidly weaker language, Spanish. The revised hierarchical model (Kroll & Stewart, 1994) and the mixed model of bilingual access (de Groot, 1992) can explain these results, as they both posit stronger connections between the conceptual system and the more proficient L1 than between the conceptual system and L2. Thus, targeting treatment at the semantic/conceptual level appears to have had a more beneficial influence on the stronger L1 than on the weaker L2. This proposal is supported by findings that the effect of phonological segments of nontarget language words appears to be related to relative proficiency of L1 and L2 (Costa et al., 1999; Roelofs, 2003). An alternate possibility may be that direct connections between L1 and L2 (bypassing the semantic system) could have been influenced by treatment. However, because connections from L1 to L2 are thought to be weaker than those from L2 to L1 in normal bilinguals with relatively high proficiency in L2 (de Groot, 1992; de Groot et al., 1994; Kroll & Stewart, 1994), this explanation seems less likely (see Figure 5).

For both P2 and P3, the presence of L2 to L1 crosslinguistic generalization and the absence of within L2 language generalization may be reconciled within the framework represented by weaker connections between conceptual memory and L2, resulting in weaker

Figure 5. Schematic of proposed semantic–lexical and lexical–lexical connections for Participants 2 and 3 to explain generalization patterns. Please see text for details.



phonological activation of nontarget phonological items for the untrained semantically related items in L2 (Spanish). The improvement in English during Spanish treatment could be due to stronger connections from L2 to L1 (de Groot, 1992; de Groot et al., 1994; Kroll & Stewart, 1994) than from L1 to L2. Thus, training language-dominant bilingual individuals in their stronger language may not result in improvements to the weaker L2, but training the weaker language may improve L1 with limited within-language generalization in L2.

Results of pre- and posttreatment standardized testing for all participants revealed improvements on tasks expected to be influenced by treatment, namely, auditory comprehension, naming, semantic processing, and reading. Semantic processing results as measured by the BAT in both languages, however, were variable and thus were inconclusive. Additionally, P2 made more improvements in English than in Spanish, thus raising the possibility that other factors in addition to the language dominance pattern could have influenced generalization. Interestingly, the WAB AQ (Kertesz, 1982) was sensitive to improvements in English for P1 and P3, even though they received treatment in Spanish only, illustrating the robustness of crosslinguistic generalization.

Extensive error analyses conducted on all participants' responses during baselines and the probe sessions revealed evolution patterns consistent with the main findings of the study. As would be expected, given the overall improvement in both languages, P1's error patterns showed increased processing in both languages with provision of treatment in only one language. Of particular note was the increase in crosslinguistic errors in both languages concomitant with the presence of crosslinguistic interference over the course of treatment for P1. This difficulty was evident in her conversational speech prior to initiation of treatment even with listeners she knew were monolingual, something she did not do premorbidly. It should be noted that the interaction between the clinician and P1 was always in the target language regardless of the language of the response produced by P1, so it is unlikely that P1's crosslinguistic interference was due to a misunderstanding of the target language.

The crosslinguistic responses for P1 can be reconciled within an existing theoretical framework explaining the mechanism by which bilingual speakers maintain the intended language. According to one view (e.g., de Bot, 1992; de Bot & Schreuder, 1993; Green, 1986, 1998; Hermans et al., 1998; Poulisse & Bongaerts, 1994), the nontarget language may not have been appropriately inhibited, and with treatment, which resulted in increased activation of the lexical item in both languages, this task became even more difficult. According to a second view (e.g., Costa et al., 1999; Costa & Caramazza, 1999; Costa & Santesteban, 2004; Roelofs, 1998), the mechanism that selects only the target language for production might be impaired and unaffected by treatment, resulting in production of either language regardless of intended language.

An increase in semantic errors over the course of treatment observed for P2 reflects an ability to activate relevant semantic information and provide a semantically related response in English, a finding consistent with previous work in monolingual aphasic individuals (Kiran & Thompson, 2003). Contrastingly, P2's Spanish errors over the course of treatment evolved into an increase in IDK/NR responses, a finding likely reflective of increased self-awareness regarding lexical retrieval in Spanish. Alternatively, it has been proposed that languages that are frequently spoken are harder to inhibit, and oftentimes are not completely deactivated when compared with languages that are not as frequently spoken (Costa & Santesteban, 2004; Green, 1986). This may have been the mechanism at work during lexical retrieval in P2 who appeared to mediate lexical access through his dominant language, and, as a result, did not show as marked a progression of errors in Spanish as he did in English. P3 predominantly produced semantic and neologistic errors during treatment, which were reduced during the course of treatment. These results are again consistent with monolingual aphasic individuals (Kiran & Thompson, 2003), thus reflecting an increasing ease of lexical access to semantic and phonological representations of targets. The different error results for the 3 participants highlight the importance of error analysis in understanding the mechanisms underlying lexical retrieval following treatment and further our current understanding of theoretical models of lexical access in bilingual individuals.

While it appears that premorbid language proficiency was the primary variable influencing the nature of within and crosslinguistic generalization, other factors warrant discussion. For instance, it appears that aphasia severity did not likely influence the amount of generalization. Whereas P2 had a more severe aphasia (with concurrent apraxia of speech) than P1 and P3, both P2 and P3 demonstrated similar generalization patterns upon training the nondominant language. It may be possible though, that P2's Spanish abilities were more severely impaired prior to treatment, but that this finding was masked by his floor-level performance on all tasks. Hence, gains were not as apparent in Spanish as in English. Second, because all participants demonstrated generalization from Spanish to English, but not vice versa (i.e., results from P2), differences across languages must be considered. Specifically, retrieval could have been affected by word length because Spanish words are typically longer than English words. To address these issues, the present results need to be replicated across languages and dominances, which we are currently pursuing.

The results of this study are preliminary due to a limited number of participants and ranges of proficiencies. However, the findings have implications for bilingual aphasia rehabilitation. Although clinically counterintuitive, it appears that training the nondominant language in an individual with bilingual aphasia may be more beneficial in facilitating crosslinguistic generalization than training the dominant language. A balanced bilingual may benefit from treatment in either language because premorbid connections may have been strong enough to allow for generalization within and across languages. However, this assertion needs further examination because the balanced bilingual in the current study was trained only in one language. Finally, more theoretically based, experimentally controlled treatment studies are crucial for guiding treatment and interpreting patterns of generalization in bilingual aphasia.

Acknowledgment

This research was supported by an American Speech-Language-Hearing Foundation New Investigator Research Grant to the second author. The authors would like to thank Ana Ayala, Keith Lebel, Lisa Rukovena, Lyena Escobar, and Zach Rogers for their assistance at various stages of the project and the 3 participants (RA, MT, and MR) for their patience and hard work throughout the duration of the study.

References

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology, 13,* 236–249.

Boyle, M., & Coehlo, C. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology, 4*(4), 94–98.

Butterworth, B. (1989). Lexical access in speech production. In W. M. Wilson (Ed.), *Lexical representation and process* (pp. 108–235). Cambridge, MA: MIT Press.

Connell, P. J., & Thompson, C. K. (1986). Flexibility of single-subject experimental designs: Part III. Using flexibility to design or modify experiments. *Journal of Speech and Hearing Disorders, 51*(3), 214–225.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language & Cognition, 2,* 231–244.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41,* 365–397.

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,* 491–511.

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language, 94,* 94–103.

de Bot, K. (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics, 13,* 1–24.

de Bot, K., & Schreuder, R. (1993). Word production and the bilingual lexicon. In R. Schreuder (Ed.), *The bilingual lexicon* (pp. 191–214). Amsterdam: John Benjamins.

de Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1001–1018.

de Groot, A. M. B. (2000). On the source and nature of semantic and conceptual knowledge. *Bilingualism, 3,* 7–9.

de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language, 33,* 600–629.

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 92,* 283–321.

Dell, G. S., Schwartz, M. F., Martin, N. M., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801–838.

Drew, R. L., & Thompson, C. K. (1999). Model-based semantic treatment for naming deficits in aphasia. *Journal of Speech, Language, and Hearing Research, 42,* 972–989.

Edmonds, L. A., & Kiran, S. (2004). Confrontation naming and semantic relatedness judgments in Spanish–English bilinguals. *Aphasiology, 18,* 567–579.

Fabbro, F. (2001). The bilingual brain: Bilingual aphasia. *Brain and Language, 79,* 201–210.

Frances, N., & Kučera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Galvez, A., & Hinckley, J. J. (2003). Transfer patterns of naming treatment in a case of bilingual aphasia. *Brain and Language, 87,* 173–174.

Goral, M., Levy, E. S., & Obler, L. K. (2002). Neurolinguistic aspects of bilingualism. *International Journal of Bilingualism, 6,* 411–440.

Green, D. W. (1986). Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language, 27,* 210–223.

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

Hamers, J. F., & Blanc, M. H. A. (2000). *Bilinguality and bilingualism* (2nd ed.). Cambridge, England: Cambridge University Press.

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, 1,* 213–229.

Hinckley, J. J. (2003). Picture naming treatment in aphasia yields greater improvement in L1. *Brain and Language, 87,* 171–172.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5,* 67–103.

Juilland, A., & Chang-Rodriguez, A. (1964). *Frequency dictionary of Spanish words*. London: Mouton.

Junque, C., Vendrell, P., Vendrell-Brucet, J. M., & Tobena, A. (1989). Differential recovery in naming in bilingual aphasics. *Brain and Language, 36*(1), 16–22.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test* (2nd ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessment of Language Processing in Aphasia*. San Antonio, TX: The Psychology Press.

Kertesz, A. (1982). *The Western Aphasia Battery.* Philadelphia: Grune and Stratton.

Kiran, S. (2005). Training phoneme to grapheme conversion for patients with written and oral production deficits: A model-based approach. *Aphasiology, 19*(1), 53–76.

Kiran, S., & Thompson, C. K. (2003). The role of semantic complexity in treatment of naming deficits: Training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research, 46,* 608–622.

Kohnert, K. J. (2004). Cognitive and cognate-based treatments for bilingual aphasia: A case study. *Brain and Language, 91,* 294–302.

Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second-language words. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory* (Vol. 2, pp. 389–395). London: Wiley.

Kroll, J. F., & de Groot, A. M. B. (2005). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford, England: Oxford University Press.

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

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Brain and Behavioral Sciences, 22,* 1–75.

Lowell, S. B., Beeson, P. M., & Holland, A. L. (1995). The efficacy of a semantic cueing procedure on naming abilities of adults with aphasia. *American Journal of Speech-Language Pathology, 4*(4), 109–114.

McReynolds, L. V., & Kearns, K. P. (1983). *Single subject experimental designs in communicative disorders*. Baltimore, MD: University Park Press.

Muñoz, M. L., Marquardt, T. P., & Copeland, G. (1999). A comparison of the code switching patterns of aphasic and neurologically normal bilingual speakers of English and Spanish. *Brain and Language, 66,* 249–274.

Paradis, M. (1987). *The assessment of bilingual aphasia*. Hillsdale, NJ: Erlbaum.

Pavlenko, A. (2000). What's in a concept? *Bilingualism, 3*(1), 31–36.

Potter, M. C., So, K., von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 23–38.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15,* 36–57.

Roelofs, A. (1998). Lemma selection without inhibition of languages in bilingual speakers. *Bilingualism, 1*(2), 94–95.

Roelofs, A. (2003). Shared phonological encoding processes and representations of languages in bilingual speakers. *Language and Cognitive Processes, 18*(2), 175–204.

Romaine, S. (1995). *Bilingualism* (2nd ed.). Oxford, England: Blackwell.

Sasanuma, S., & Suk Park, H. (1995). Patterns of language deficits in two Korean–Japanese bilingual aphasic patients—A clinical report. In M. Paradis (Ed.), *Aspects of bilingual aphasia* (pp. 111–122). Oxford, England: Pergamon.

Tryon, W. W. (1982). A simplified time-series analysis for evaluating treatment interventions. *Journal of Applied Behavior Analysis, 15,* 423–429.

U.S. Bureau of the Census. (2000). *Census report 2000*. Washington, DC: Author.

Watamori, T. S., & Sasanuma, S. (1978). The recovery processes of two English–Japanese bilingual aphasics. *Brain and Language, 6,* 127–140.

Received March 21, 2005

Revision received August 1, 2005

Accepted November 29, 2005

DOI: 10.1044/1092-4388(2006/053)

Contact author: Swathi Kiran, CMA 7.206, Department of Communication Sciences and Disorders, University of Texas, Austin, TX 78712. E-mail: s-kiran@mail.utexas.edu

Lisa A. Edmonds is now at the Department of Communication Sciences and Disorders, University of Florida.

Appendix. Average frequencies for treatment stimuli for the 3 participants.

	English Set 1	Spanish Set 1	English Set 2	Spanish Set 2	English UR	Spanish UR
			Participant 1			
	Celery	Apio	Cabbage	Repollo	Trunk	Bombero
	Dustpan	Recogedor	Vacuum (cleaner)	Aspiradora	Clothespin	Secadora (de pelo)
	Wrench	(Llave) perica	Drill	Taladro	Pitcher	Sobre
	Snail	Caracol	Worm	Gusano	Owl	Zancos
	Skunk	Zorrillo	Raccoon	Mapache	Goat	Ganso
	Shelf	Estante	Hook	Gancho		
	Robe	Bata	Coat	Abrigo		
	Wheelbarrow	Carretilla	Rake	Rastrillo		
	Forehead	Frente	Chin	Mentón		
	Barn	Granero	Skyscraper	Rascacielos		
Average (*SD*)	7.3 (10.6)	2.0 (4.9)	10.6 (16.5)	4.6 (13.8)		
			Participant 2			
	Orange	Naranja	Apple	Manzana	Star	Vaca
	Table	Mesa	Chair	Silla	Skirt	Taza
	Fork	Tenedor	Spoon	Cuchara	Shovel	Barco
	Fish	Pez	Shark	Tiburón	Ball	Cobija
	Ring	Anillo	Necklace	Collar	Clown	Libro
	Hat	Sombrero	Gloves	Guantes		
	Door	Puerta	Window	Ventana		
	Arm	Brazo	Leg	Pierna		
	Garlic	Ajo	Onion	Cebolla		
	Spider	Araña	Ant	Hormiga		
Average (*SD*)	82.3 (117)	96.6 (169.4)	51.1 (62.18)	63.9 (78.2)		
			Participant 3			
	Ant	Hormiga	Spider	Araña	Apron	Escalera
	Orange	Naranja	Apple	Manzana	Wheelbarrow	Escritorio
	Cabbage	Repollo	Radish	Rábano	Ear	Bata
	Purse	Bolsa	Hat	Sombrero	Necklace	(Llave) perica
	Sheep	Borrego	Donkey	Burro	Pliers	Pulsera
	Newspaper	Periódico	Magazine	Revista		
	Eagle	Águila	Owl	Búho		
	Raccoon	Mapache	Skunk	Zorrillo		
	Shark	Tiburón	Whale	Ballena		
	Snail	Caracol	Worm	Gusano		
Average (*SD*)	19.2 (30.8)	23.5 (41.3)	14.2 (23.2)	23.5 (32.4)		

Note. UR = semantically unrelated (control set).