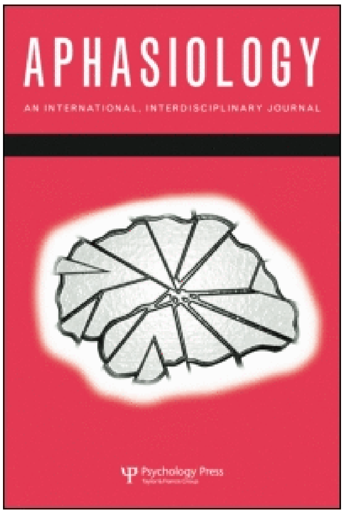


## Aphasiology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713393920

# Semantic feature analysis treatment in Spanish-English and French-English bilingual aphasia

Swathi Kiran [a]; Patricia M. Roberts [b]
[a] Boston University, Boston, MA, USA [b] University of Ottawa, Canada

First published on: 31 August 2009

APHASIOLOGY
2009, iFirst Article, 1–31

Psychology Press
Taylor & Francis Group

# Semantic feature analysis treatment in Spanish–English and French–English bilingual aphasia

Swathi Kiran

*Boston University, Boston, MA, USA*

Patricia M. Roberts

*University of Ottawa, Canada*



*Background:* Edmonds and Kiran (2006) reported that training lexical retrieval in one language resulted in within-language and cross-language generalisation in three bilingual (English–Spanish) patients with aphasia.

*Aims*: The present experiment continues this line of research, repeating a similar procedure with new patients and examining a broader range of factors that may affect generalisation patterns.

*Methods & Procedures*: Four participants (two Spanish–English and two French–English speakers) with anomia post CVA received a semantic feature-based treatment aimed at improving naming of English or Spanish/French nouns. Using a multiple baseline design, generalisation to untrained semantically related and unrelated items in each language was measured during periods of therapy first in one language, then in the other.

*Outcomes & Results*: All patients improved their naming of the trained items in the trained language, although to varying degrees. Within-language generalisation to semantically related items occurred in two Spanish–English patients and one French–English patient. Cross-language generalisation to translations and semantically related items occurred only for one French–English patient.

*Conclusions*: The impact of the intervention is very clear. The semantic feature-based practice is linked to the gains made, and accounts for the predominance of semantic naming errors after treatment. Possible explanations for the different patterns of generalisation are considered in terms of the various factors including each patient's pre-stroke language proficiency, age of acquisition of each language, post-stroke level of language impairment, and type and severity of aphasia.

***Keywords:*** Cross-language generalisation; Bilingual aphasia; Naming treatment; Semantic feature analysis; Aphasia.

Address correspondence to: Swathi Kiran, E-mail: kirans@bu.edu or Pat Roberts, proberts@uottawa.ca

This research was supported by an American Speech and Hearing Foundation New Investigator Grant to the first author. The authors thank Claudine Choquette, Mélanie Cohen, Anne Duff, and Deborah Kauffman for their work with the four patients. The translation advice of Linguistek and the University of Ottawa translation service was also very helpful. And, above all, we thank the four volunteers with aphasia who participated in the study over many hours of repetition and testing. Both authors contributed equally to this paper.


http://www.psypress.com/aphasiology DOI: 10.1080/02687030902958365

As clinicians face increasing numbers of bilingual patients and limited time for their treatment, the question of generalisation of treatment effects across languages in bilingual patients is an important one in aphasiology. On a clinical level, understanding why treatment in one language sometimes improves an untreated language, to varying degrees, and why it sometimes has little to no effect on the untreated language will help clinicians achieve maximum benefit for their patients in the minimum amount of time. To the extent that treatment effects generalise across languages, clinicians who do not speak the languages of their patients can nonetheless expect gains in these other languages following treatment in only one of the patient's languages (Roberts, 1998). To the extent that treatment gains are limited to the language of treatment, patients must receive treatment in all languages they need to use. On a theoretical level, studies of the effects of bilingual aphasia treatment can shed light on the structure and functioning of the bilingual language system and contribute evidence for or against various models of bilingual language processing.

Studies of cross-language generalisation have yielded a range of results. At times, the treated language improves more than the untreated one (i.e., limited cross-language generalisation), while other studies have found other patterns (see Roberts & Kiran, 2007, for a recent review of this literature).

One of the drawbacks of treatment case studies in bilingual aphasia is that they have only one or two patients and only one pair of languages in each study. As noted by Roberts and Kiran (2007), the variability inherent in both aphasia and bilingualism makes it difficult to interpret results obtained in therapy. Before drawing firm conclusions about the effects of various treatments and their implications for models of bilingualism, results must be shown to be consistent across patients and across studies. The primary aim of the present study is therefore to replicate and extend the work of Edmonds and Kiran (2006), by examining the results of treatment for anomia in new patients, including a new pair of languages (French–English).

Studies of normal bilingual individuals offer tests of models of lexical semantic representations and processing (de Groot, 1992; Dijkstra & Van Heuven, 2002; Green, 1998; Kroll & Stewart, 1994; Potter, So, von Eckardt, & Feldman, 1984). These models generally agree that bilingual individuals have a shared semantic/conceptual system and that there are separate lexical representations for the two languages. Some studies posit that the semantic system spreads its activation to lexical items in both languages regardless of the target language (for a recent review see Costa, La Heij, & Navarrete, 2006) with the size and direction of this spreading activation depending, to some extent, on age of acquisition and proficiency levels in each language (e.g., Silverberg & Samuel, 2004) as well as on the semantic similarity between the specific lexical items. To the—currently unknown—extent that this is true for patients with damaged language systems, generalisation of gains could occur across languages when semantic representations in one language are repeatedly activated during therapy.

For monolingual patients with aphasia, treatments based on models of lexical semantic processing that emphasise strengthening semantic information at the semantic/conceptual level have been successful in facilitating generalisation to untrained semantically related items in some patients (Coelho, McHugh, & Boyle, 2000; Drew & Thompson, 1999; Kiran & Thompson, 2003). Therefore, training that strengthens or emphasises semantic features should increase the level of activation of trained items and of their semantically related neighbours, thereby facilitating generalisation to untrained semantically related items. It is also hypothesised that 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 in both lexicons (e.g., Costa & Caramazza, 1999). Hence, the lexical forms that receive treatment in the target language will also activate translation equivalents in the non-target language, leading to improvement in these items as well.

In order to test the predictions outlined above, Edmonds and Kiran (2006) administered a semantic treatment to improve picture naming in English and Spanish and measured generalisation to translations of the treated words and to words semantically related to the target words, in each language. Results demonstrated within- and across-languages effects on generalisation that the authors interpreted as being related to pre-morbid language proficiencies. Participant 1, who claimed equal proficiency in English and Spanish, showed within-language generalisation in the trained language (Spanish) and some cross-language generalisation to the untrained language (English). The other two patients, who reported that English was their stronger language pre-stroke, showed cross-language generalisation from the trained language (Spanish) to the untrained language (English) but no within-language generalisation (to related words in Spanish). The results could be interpreted as showing that training the patient's weaker language may facilitate more cross-language generalisation than training their stronger language, but more patients need be examined to explore this very preliminary hypothesis.

One problem with interpreting the Edmonds and Kiran (2006) results in terms of reported pre-stroke level of bilingualism is the unknown validity of self-reports of proficiency. For example, their patient P1 claimed equal proficiency in English and Spanish for auditory comprehension, reading, and writing, and only slightly better Spanish for verbal expression. This profile is very surprising in someone who also stated that "she acquired English as an adult" (p. 732).

There are other factors that may influence the extent of cross-language generalisation. These include factors always present in studies of aphasia: the type and severity of the aphasia, lesion location (both macro and micro), and degree of impairment in each language. Additionally, factors related to bilingualism may be relevant: including domains of language use, age of acquisition, and the particular linguistic features of Spanish–English bilingualism. As noted by the authors, one limitation in the Edmonds and Kiran (2006) study was that only one of the three patients received therapy in both languages.

The present study examines the effects of a primarily semantic treatment on anomia in four patients and two language pairs (Spanish–English and French–English), exploring the same hypotheses as Edmonds and Kiran (2006). The theoretical foundations for these hypotheses are presented in detail in the 2006 paper. It was hypothesised that a primarily semantically based treatment would lead to:

1. gains on *trained items* in the *trained language* (e.g., English: *door*);
2. within-language generalisation to semantically related, not treated items (i.e., if *door* is practised in treatment, we expect improvement in *window*), if the gains in 1 are seen;
3. cross-language generalisation to the translations of the treated items (i.e., if *door* is practised in treatment, we expect improvement in *puerta* or in *porte*);
4. cross-language generalisation to items semantically related to the treated items in the untrained language (i.e., if *door* is practised, we expect improvement in the Spanish or French words for window: *ventana* or *fenêtre*).
5. No changes in semantically unrelated, "control words" in either language.

# METHOD

## Participants

Four bilingual women with aphasia (two English–Spanish and two French–English) were recruited from local area hospitals and stroke support groups in Austin, Texas, USA and Ottawa, Ontario, Canada. All participants met the following selection criteria: (a) diagnosis by a neurologist of a single stroke in the left hemisphere (encompassing the grey/white matter in and around the perisylvian area) confirmed by a CT/MRI scan; (b) onset of stroke at least 6 months prior to participation in the study; (c) right-handed prior to stroke; (d) adequate hearing, vision, and comprehension to engage fully in testing and treatment; (e) stable health status; (f) previous speech-language therapy had ended at least 1 month prior to starting the present study; (g) bilingual speakers of English and Spanish or English and French who reported regular use of both languages prior to their stroke; and (h) not able to speak any other language. See Table 1 for demographic details. None of the participants had diabetes. None reported any cognitive or communication problems prior to their stroke. All except P2 held full-time jobs prior to their stroke.

Language history, use patterns (Muñoz, Marquardt, & Copeland, 1999; Roberts & Shenker, 2007), and proficiency ratings were used to estimate pre-morbid proficiency in both languages. In addition to self-ratings, one bilingual family member familiar with the participant's language acquisition and use was also interviewed to corroborate information provided by the participants.

### *Participant 1 (P1)*

Participant 1 had a single left hemisphere CVA 11 months prior to starting this study, and was diagnosed with severe expressive aphasia. Most of the therapy she received was in English. The patient reported that after her stroke she had more difficulty communicating in Spanish than in English. P1 presented with hemiparesis of her right hand and was diagnosed with verbal apraxia by the referring speech pathologist.

*Language background.* Participant 1 grew up in a predominantly Spanish-speaking home in Texas. She probably had some passive exposure to English before starting an English-language elementary school at age 5. As an adult she spoke in English and Spanish to her spouse, her children, and all relatives. However, she rated herself as being more proficient in English than in Spanish in her daily interactions with friends and family members. P1 spoke English with no Spanish accent.

*Education and work history.* P1 was educated in English and taught herself to read and write in Spanish up to high-school level. She worked as a clerk in a community education classroom for English as a second language. Therefore, prior to her stroke, she reported that she used English and Spanish relatively equally at work.

### *Participant 2 (P2)*

Participant 2 had a single left hemisphere stroke approximately 6 months prior to the present study. She was diagnosed with receptive aphasia after her stroke and received therapy mainly in English. Although P2's aphasia resolved considerably in the first 6 months, she still presented with comprehension difficulties. She reported that the stroke affected both languages equally.

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

| *Demographic information* | | | | | | *Language history and proficiency* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Pt* | *Sex* | *Age* | *Education* | *Aetiology* | *MPO* | *Family/social* | *Work* | *Reading/writing* | *Self-ratings (L1/L2) (1–7)* | *BPR* |
| 1 | F | 55 | 14 yrs? | Left MCA | 11 | Born in US.<br>Began English at age 5.<br>Spanish from birth.<br>Married to bilingual Spanish speaker.<br>Spoke both English and Spanish with children, siblings, with friends. | English: 50%<br>Spanish 50%<br>Clerk in community education for English as a second language | Educated in English<br>Self-taught Spanish<br>Read and wrote English and Spanish materials | Speech 6/7<br>Comp 6/7<br>Reading 4/7<br>Writing 4/7 | .78<br>(English dominant) |
| 2 | F | 87 | 12 years | Left MCA | 6 | Spanish only with parents, siblings, relatives, friends.<br>English with grandchildren and other professionals. | 50% English<br>50% Spanish<br>Writer of Mexican fiction books | Educated in Spanish<br>Wrote letters and lists in Spanish<br>Learned and used English<br>Read and wrote English and Spanish materials | Speak: 7/7<br>Comp: 7/7<br>Read: 7/7<br>Write: 7/7 | 1.0<br>(Equally proficient) |
| 3 | F | 55 | 12 years | Left MCA | 33 | Grew up in bilingual environment but much more exposure to French than English. French spoken at home as child and mostly French as adult. | 70% French; 30% English. Bank clerk, secretary; teacher's aide. | Education in French, except for 1 year of community college.<br>Strong preference for French. Reads a little for pleasure, mostly in French | Speech 6/4<br>Comp 7/5<br>Reading 6/4<br>Writing 6/4 | 1.4<br>(French dominant) |

(*Continued*)

TABLE 1
Continued

| *Demographic information* | | | | | | *Language history and proficiency* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Pt* | *Sex* | *Age* | *Education* | *Aetiology* | *MPO* | *Family/social* | *Work* | *Reading/writing* | *Self-ratings (L1/L2) (1–7)* | *BPR* |
| 4 | F | 60 | 15 years | Left MCA | 15 | Grew up in French environment. Learned English from grade 3 onwards. Adult social life mostly in French | 80% English 20% French; federal civil servant with administrative and training roles | Schooling primarily in French except university, which was in both languages. Avid reader of both languages with more exposure to English. | Speech 7/6<br>Comp 7/6<br>Reading 7/7<br>Writing 7/6 | 1.1 (Equally proficient/ French dominant) |

MPO: Months Post Onset; E: English; S: Spanish; Comp: Comprehension; MCA: Middle Cerebral Artery; CVA: Cerebral Vascular Accident; BPR: Bilingual Proficiency Ratio.

*Language background.* P2 grew up in a predominantly Spanish-speaking environment in Mexico. Her exact age of immigration to the US is not known. As a child she spoke only Spanish with her parents, siblings, and relatives. As an adult she spoke English with most of her grandchildren and friends and other professionals. She reported that her linguistic environment as an adult was "100% bilingual" although she was more comfortable in Spanish than English. She spoke English with a strong Spanish accent.

*Education and work history.* P2 was educated in Spanish and took classes in English when she arrived in the US. Specific details about her education history are not available. She was a successful author, writing non-fiction books in Spanish and also worked in a Mexican art museum.

### *Participant 3(P3)*

Participant 3 had a single left hemisphere CVA approximately 35 months prior to starting this study. She initially experienced quite severe aphasia and received therapy in French for approximately 3 months following her stroke. The aphasia resolved into a relatively mild aphasia, which seemed to affect both languages equally. French was her preferred/stronger language pre-stroke and remained so post-stroke. P3's initial hemiplegia had resolved into hemiparesis. However, she remained unable to use her right hand to write.

*Language background.* P3 reported that her family spoke French 90% of the time at home when she was a child. She reported some exposure to English at home, and through television prior to that, but the small rural Ontario community she grew up in was primarily French. She spoke English with a noticeable French accent, somewhat stronger than that of P4.

*Education and work history.* Language of instruction at school was French. P3 studied English as a subject beginning around age 10 or 11. She finished high school and took some courses at community college level but did not complete a diploma. She worked at various white-collar jobs, including bank teller (primarily in English) and, for the 10 years prior to her stroke, as a teaching assistant (primarily in French).

### *Participant 4 (P4)*

Participant 4 experienced a left hemisphere stroke 14 months prior to starting this study. Initially she was diagnosed with moderate non-fluent aphasia with agrammatism, moderately severe verbal apraxia, and mild to moderate dysphagia. She received language therapy in English. Prior to beginning this study her aphasia was mild. Mild verbal apraxia remained, giving her speech a slightly slow and effortful quality and occasionally interfering with intelligibility. A mild French accent was present when she spoke English. Prior to this study she received individual speech-language therapy as an inpatient, then as an outpatient, primarily in English, with the most recent (group) treatment several months before the start of this study. She felt the stroke had affected the two languages about equally.

*Language background.* P4's family spoke French when she was a child in rural Ontario, near Ottawa. She reports beginning English classes in school around age 8 but had some exposure to English prior to age 8 through friends and neighbours. The language of instruction was French for elementary and high school, while university was in both languages.

*Education and work history.* P4 completed a bachelor's degree in psychology. At university, readings and classes were in both French and English. She worked at various white-collar jobs, primarily in the federal civil service. Prior to the stroke she had held an administrative and training position for many years where she used English 80% of the time and French 20% of the time.

To quantify pre-morbid language proficiencies, and to facilitate comparisons across patients, a bilingual proficiency ratio (BPR) was calculated by dividing the sum of the self-ratings in Spanish or French by the sum of the self-ratings in English: BPR=(Spanish/French comprehension+verbal expression+reading+writing)/(English comprehension+verbal expression+reading+writing). A BPR of 1 reflects equal language proficiency. Scores less than 1.0 reflect dominance in English, whereas scores greater than 1.0 reflect dominance in Spanish or French. Therefore, based on their self-reports, P1 was relatively English dominant (.78), P2 was equally proficient (1.0), P3 was relatively more dominant in French (1.4) and P4 was very close to balanced, but with a slight preference for French (1.1).



### Language status after stroke

Prior to starting treatment, participants completed relevant portions of the *Western Aphasia Battery* (WAB; Kertesz, 1982) in English, several semantic and lexical subtests from the *Psycholinguistic Assessment of Language Processing in Aphasia* (PALPA; Kay, Lesser, & Coltheart, 1992) in English, the *Boston Naming Test* (BNT; Kaplan, Goodglass, & Weintraub, 2001) in both languages, and portions of the *Bilingual Aphasia Test* (BAT; Paradis, 1989) in both languages. Scores for the WAB, BNT, and the BAT cannot be interpreted as showing the relative severity of the aphasia in each language. However, they were administered to allow comparisons pre/post-therapy for each participant on each test. Summaries for each participant are provided in Tables 2 and 3.

Participant 1 was diagnosed with moderate aphasia characterised by nonfluent speech, relatively intact comprehension, and mild difficulties in repetition and naming based on the *WAB*. Performance on the two word-to-picture matching subtests of the PALPA revealed mild impairments. Her score on the BNT was 6/60 in Spanish and 34/60 in English. Performance on the various subtests of the BAT revealed an overall superior performance in English across receptive tests such as complex commands, judgement of words/nonwords, and semantic categories. The stronger performance in English was marked on tests that required verbal output such as naming, sentence construction, antonym distinction, reading text, and dictation. P1's comprehension in Spanish was much better than her production abilities.

Participant 2 had a moderate to severe aphasia characterised by fluent speech, impaired comprehension, and poor repetition and naming. Performance on the two word-to-picture matching subtests of the PALPA revealed moderate impairments. Performance on the BNT revealed severe naming deficits in both languages (5/60 in English and 3/60 in Spanish). Performance on the various subtests of the BAT revealed similar performance in English and Spanish on receptive tests such as pointing, semicomplex commands, grammaticality judgement, and judgement of

TABLE 2
Pre- and post-language performance on tests administered in English only: WAB (Kertesz, 1982) and PALPA (Kay et al., 1992)

| | *Participant 1* | | *Participant 2* | | *Participant 3* | | *Participant 4* | |
|---|---|---|---|---|---|---|---|---|
| *Test* | *Pre-tx* | *Post-tx* | *Pre-tx* | *Post-tx* | *Pre-tx* | *Post-tx* | *Pre-tx* | *Post-tx* |
| *Western Aphasia Battery (WAB)* | | | | | | | | |
| Spontaneous speech | 11 | 12 | 14 | 16 | 15 | 15 | 16 | 16 |
| Auditory comprehension | 9.2 | 8.9 | 4.1 | 5.6 | 9.8 | NA | 8.9 | NA |
| Repetition | 7.1 | 6.4 | 2.3 | 2.9 | 9.5 | NA | 8.6 | NA |
| Naming | 7.6 | 7.7 | 4.0 | 5.4 | 7.2 | 7.9 | 8.2 | NA |
| *PALPA* | | | | | | | | |
| Spoken Word–Picture Matching (%) | 90 | 90 | 80 | 70 | 90 | NA | 98 | NA |
| Written Word–Picture Matching (%) | 90 | 92.5 | 87.5 | 90 | 98 | NA | 98 | NA |
| Auditory Synonym Judgements (%) | 72 | 75 | 48 | 45 | NA | NA | NA | NA |
| Written Synonym Judgements (%) | 85 | 77 | 90 | 66.7 | 87 | NA | NA | NA |

NA indicates not administered.

words/nonwords. However, performance in English was superior to Spanish on subtests involving reading and verbal output.

Participant 3 displayed a mild anomic aphasia, with auditory comprehension superior to verbal expression, good repetition abilities, and mild to moderate naming deficits in both languages. She obtained near perfect scores on the PALPA subtests examining semantic processing. BNT scores were lower in English (22/60) than in French (39/60). On the BAT she showed some difficulty with syntactic structures. P3's performance was generally high in both languages, however relatively superior performance was observed for French on semantic categories, judgement of real words/nonwords, and semantic acceptability.

Participant 4 displayed a mild nonfluent aphasia characterised by relatively intact auditory and reading comprehension, good repetition abilities, mild agrammatic tendencies in spontaneous speech, and mild to moderate naming deficits in both languages. Performance on the two subtests of the PALPA revealed relatively normal performance. Performance on the BNT revealed mild naming impairments in English (42/60) and in French (43/60). Likewise she displayed mild deficits on various subtests of the BAT that included judgement of real/non words, repetition, semantic categories, semantic antonyms, synonyms, naming, and reading. In verbal fluency she performed better on semantic categories than on phonological ones. Overall, this patient demonstrated milder impairments than the three other participants, in both languages.

## Stimuli

Five stimulus sets were created for each patient. The first set (English Set 1) consisted of picturable English nouns. French or Spanish Set 1 was the translation of English

TABLE 3
Performance selected sub-tests of Bilingual Aphasia Test before and after treatment

| | *P1* | | | | *P2* | | | | *P3* | | | | *P4* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *PRE-TX* | | *POST-TX* | | *PRE-TX* | | *POST-TX* | | *PRE-TX* | | *POST-TX* | | *PRE-TX* | | *POST-TX* | |
| | *E* | *S* | *E* | *S* | *E* | *S* | *E* | *S* | *E* | *F* | *E* | *F* | *E* | *F* | *E* | *F* |
| Task | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* | *%* |
| Semi-complex Commands (10 points) | 90 | 70 | 100 | 60 | 50 | 60 | 30 | 60 | 100 | 100 | NA | NA | NA | NA | NA | NA |
| Complex Commands (20 points) | 100 | 55 | 55 | 60 | 15 | 5 | 25 | 5 | NA | 80 | NA | NA | 85 | NA | NA | NA |
| Semantic Categories (5 points) | 80 | 40 | 80 | 80 | 20 | 0 | 60 | 40 | 80 | 100 | NA | 80 | 80 | 80 | 80 | 80 |
| Synonyms (5 points) | 100 | 80 | 100 | 60 | 20 | 0 | 0 | 0 | 100 | 100 | NA | 100 | 100 | 80 | NA | NA |
| Antonyms I (5 points) | 100 | 60 | 100 | 40 | 20 | 0 | 20 | 0 | 80 | 80 | NA | 60 | 80 | 80 | 80 | 100 |
| Antonyms II (5 points) | 40 | 40 | 40 | 40 | 40 | 40 | 20 | 40 | 80 | 60 | NA | 60 | 80 | 80 | 80 | 80 |
| Grammaticality Judgement (10 points) | 50 | 70 | 70 | 50 | 50 | 60 | 50 | 40 | 80 | 100 | NA | NA | 80 | 80 | 90 | 80 |
| Semantic Acceptability (10 points) | 40 | 30 | 100 | 80 | 40 | 50 | 80 | 60 | 90 | 90 | NA | NA | 90 | 90 | 90 | 80 |
| Repetition (30 points) | 77 | 67 | 77 | 87 | 13 | 37 | 37 | 37 | 93 | 100 | NA | NA | 93 | 100 | 90 | 97 |
| Judgment of words/nonwords (30 points) | 90 | 73 | 97 | 90 | 70 | 60 | 67 | 73 | 87 | 100 | NA | NA | 87 | 93 | 80 | 97 |
| Series (automatics) (3 points) | 33 | NA | 0 | 0 | 0 | 0 | 33 | 67 | NA | 100 | NA | NA | 100 | 100 | NA | NA |
| Naming (19 points) | 100 | 26 | 100 | 32 | 25 | 25 | 50 | 25 | NA | 100 | NA | NA | 90 | 100 | NA | NA |
| Semantic Opposites (10 points) | 40 | NA | 30 | 10 | 0 | 0 | 60 | 10 | NA | 100 | NA | NA | NA | 80 | 70 | 90 |
| Reading Words (10 points) | 90 | 80 | 100 | 80 | 90 | 40 | 100 | 100 | NA | 100 | NA | NA | 100 | 100 | NA | NA |

E = English, S = Spanish, F = French.

Set 1. For each item in English Set 1 a semantically related category coordinate was included in English Set 2. French/Spanish Set 2 was a translation of English Set 2. Unrelated words made up Set 3, a control set. All words were non-cognates (defined as <50% overlap in phonemes). For P1 and P2 it was possible to find 15 words for Sets 1 and 2, providing a larger data set with which to measure changes during treatment. For P3 and P4 this was not feasible, so lists of 10 items were used, as in the Edmonds and Kiran study. See the appendix for lists of the stimuli used with each patient.

We attempted to match for perceived and written word frequency and for word length across the five lists, but the primary criterion was the patient's inability to name these pictures during baseline testing. The English words were verified in Francis and Kucera (1982), Thorndike and Lorge, and Brown at the MRC database site (http://www.psych.rl.ac.uk/MRC_Psych_Db.html). Spanish word frequency values were obtained from published norms (Juilland & Chang-Rodriguez, 1964). None of the published databases consulted contained word frequencies for all French words: Baudot (1992); Frantext (http://www.frantext.fr/noncateg.htm); Lexique 3 (www.lexique.org). Furthermore, homonyms/homographs used for all four participants were not listed separately (e.g., *bit*, *bolt*, *button*, *leg*, *counter*, *mouth* in English; v*olet*, and *macaron*, in French). Therefore, valid word frequency values for all stimuli could not be obtained.

Ratings of semantic relatedness for the stimuli for P1 and P2 were based on previous studies (Edmonds & Kiran, 2006; Kiran & Lebel, 2007), using a 4-point scale. For P1 and P2, the mean ratings for semantic related pairs were 2.2 and 2.2 respectively. Ratings for semantic relatedness for stimuli for P3 and P4 were obtained from 19 naive judges on a 5-point scale. The means were 2.2 for P3 and 2.0 for P4. See the appendix for the stimulus lists. Colour pictures were chosen from Art Explosion Software® (NOVA Inc), and from C-O-L-O-U-R library photos from Communication Skill Builders. For *clasp*, *button*, and *glasses case*, the actual objects were used.

### *Development of semantic features for treatment*

As in Edmonds and Kiran (2006), we developed a set of 12 yes/no questions for each target word similar to the approach of semantic feature analysis treatment, (Boyle, 2004; Boyle & Coehlo, 1995; Coelho et al., 2000). For six questions the answer was "yes" and for six it was "no". Questions for each item focused on: (1) the superordinate category (e.g., *fruit*; *insect*); (2) function or common use; (3) general characteristic "is a" (e.g., *is sweet*; *is made of metal*); (4) physical characteristic (e.g., *has skin/peel*; *has wings*); (5) typical location; (6) a personal association for each patient (*reminds me of* …). The associations were worked out with each patient in the first one or two sessions, with assistance from the clinician as needed.

## Design

The experimental design was single-participant experimental across multiple behaviours and participants. All participants received three initial baseline sessions or enough testing sessions to find the required number of stimuli that the patient was consistently unable to name, in both languages. Treatment then began on one set of items in one language (e.g., English set 1). The remaining three sets of stimuli

(English set 2, French set 1, and French set 2) and the control items were verified after every second treatment session. For details about the order of treatment, see Table 4. The language of the first block of treatment was determined prior to recruiting the participants and was counterbalanced across participants. The goal was for patients to reach 80% accuracy in naming pictures in two of three consecutive sessions. If this level was not reached after 20 treatment sessions, treatment on that set of targets stopped. Subsequently, treatment was shifted to the semantically related set in the untrained language (e.g., if *shark* was trained in English then *ballena* was trained in Spanish).

## Baseline measures

Baseline testing took place over several sessions, alternating which language was tested first and varying the order of stimuli each time. For the baseline and treatment probes, responses were considered correct if they were the expected (standard) name or a regional or slang term in wide use. Self-corrected responses were scored as correct. Minor apraxic or dysarthric errors were disregarded, except where a substituted phoneme resulted in a different word (e.g., *dog/bog*). All responses were coded into one of 10 categories. These were: (a) no response, (b) neologism or perseveration (defined as a repetition of a word at least three times within the same session), (c) unrelated word with no semantic or phonemic relationship to the target word, (d) phonemic error in target language (e.g., English – *pader* for *spider*), (e) semantic error in target language (e.g., English – *cabbage* for *radish*), (f) circumlocution, (g) mixed, semantic, and phonemic error (e.g., English *pur* for *wallet*), (h) phonemic error in nontarget language (e.g., *hooka* for *gancho* in Spanish); in some cases this error type corresponds to what other authors have called "false cognates" (e.g., Roberts & Deslauriers, 1999), (i) semantic error in nontarget language (e.g., *mesa* in Spanish for *chair*) and (j) correct response in nontarget language (e.g., *gato* in Spanish for *cat*).

For P1 it was difficult to settle on a set of items that were impaired in English and in Spanish, for two reasons. The patient was much more impaired in Spanish than in English naming, and her naming scores in English improved during baseline testing. Therefore after two attempts at finding stimuli she could not name in both English and Spanish, the focus shifted to finding items she was unable to name in Spanish. Consequently, P1's naming accuracy during baseline is higher in English than in

TABLE 4
Order of language and stimulus trained for each participant in the study

| *Participant* | *Language* | *Stimulus set trained* |
|---|---|---|
| P1 | 1. Spanish | Spanish set 1 |
| P2 | 1. English<br>2. Spanish | English set 1<br>Spanish set 2 |
| P3 | 1. French<br>2. English | French set 1<br>English set 2 |
| P4 | 1. English<br>2. French | English set 1<br>French set 2 |

Spanish. Both P3 and P4 tended to improve during the baseline phase, which meant it was difficult to find sets of non-cognate words that they were consistently unable to name in both languages and for which clear pictures and a suitable semantically related word, also consistently failed, in both languages were available. Some words were eliminated as potential stimuli if the patient indicated that she did not recognise and/or would not have known the name pre stroke. Both women reported that sometimes the word for a picture they were unable to name during testing suddenly "came back" after testing, allowing them to correctly name that item in the next session. Sometimes this occurred before leaving the testing room. They also asked family members to provide names between sessions. For these reasons, the stimuli for each French–English participant included one pair that had weaker semantic relatedness than the preset minimum of 2.5/5 and three pairs that were closer to being cognates than optimal.

## Treatment

Each set of target items was practised using a seven-step semantic feature analysis treatment method (Boyle & Coehlo, 1995; Edmonds & Kiran, 2006; Haarbauer-Krupa, Moser, Smith, Sullivan, & Szekeres, 1985) with the following characteristics: (1) the patient attempted to name the picture and was told if their answer was correct or not; (2) the clinician named the object; (3) the clinician placed the printed name of the picture on or below the picture; (4) the patient read a short sentence or phrase describing one of 12 semantic features of the object; (5) then sorted them into piles/groups of correct/incorrect features; (6) the picture was turned over for P1 and P2, but was left visible for P3 and P4 and the participants were asked the same 12 yes/no questions regarding these features (e.g., *Is it a fruit? Is it found on the roof?*); and (7) the patient named the picture again. Even if the participant named the picture correctly in Step 1, the whole procedure was followed. For the first sessions the clinician explained what each question was about: category name, function, etc., to make patients aware of the structure of the therapy.

### *Treatment probes*

Every second session began with the administration of probes to measure progress on the target words and generalisation within and across languages to the related words and the control set of words. The language first assessed alternated from one probe session to the next. Responses to naming probes are the primary dependent measure for P1 and P2. For P3 and P4, because of the rapid improvement, their scores during treatment sessions are the dependent measures during the treatment phase in each language.

## Data analysis

Effect sizes (ES) were calculated comparing the mean of all data points in the treatment and maintenance phases to the baseline mean divided by the standard deviation of baseline (Beeson & Robey, 2006; Busk & Serlin, 1992). When there was no variation in the baseline (e.g., accuracy was at 0% across the three baselines), the zero variance value was replaced with the mean variance of other baseline phase data for the same individual as recommended by Beeson and Robey (2006). Data from the follow-up phase were not included in the analysis. When treatment was shifted to the second language, the data preceding the onset of treatment were included within the baseline phase for that condition for that participant. Finally, because the French–English treatment comprised fewer sessions than the Spanish–English treatment, effect sizes for the generalisation (non-trained and control items) are calculated over the entire period and not specific to each treatment phase. Based on comparable naming treatment studies in aphasia, an ES of 4.0 was considered small, 7.0 was considered medium, and 10.1 was considered large (Beeson & Robey, 2006).

## Reliability

All baseline and probe sessions were recorded on audiotape and/or videotape and scored by two different individuals. Point-to-point agreement was ⩾90% across all sessions. Discrepancies were resolved by repeated listening to the tapes. Scoring of regional terms and decisions about acceptable synonyms were made after consulting dictionaries, other bilingual speakers, and professional translators.

# RESULTS

## Naming accuracy

Results for each patient are displayed in Figures 1 to 4. For reasons discussed above, data for P3 and P4 are their naming scores during the treatment phase. For P1 and P2, data are their naming scores during the weekly probes.

### *Participant 1*

P1 was trained in Spanish (her weaker language). After three baseline sessions, naming of set 1 improved to a high of 53%, which did not meet criterion despite the effect size of 7.4. Concomitant changes were observed in the semantically related but untrained Spanish set 2 (ES = 1.2), which improved to a high of 47%. English translations (English set 1) of the trained items also improved to a high of 70% (ES = 2.2), continuing the rising baselines pattern pre-therapy. Performance on untrained, English, semantically related items (English set 2) showed a slight improvement but performance fluctuated between 40 and 67% (ES = 2.3). Because performance on both English sets improved during Spanish treatment to reach over 60% accuracy, treatment was not provided in English.

Performance on the English unrelated control items showed an improvement from 33% to a high of 83% (ES = 3.5). Although this change appears to be quite dramatic it is amplified because there were only five items in this set. No changes were observed in the Spanish unrelated control words (ES = 0.2).

### *Participant 2*

*P2* was trained first in English and then in Spanish. On English set 1 performance improved to 66% (ES = 10.1) after nine sessions. Performance on the semantically related items in English (English set 2) showed a trend of improvement (from 6% to a maximum score of 40%). Performance on the Spanish translations of the items trained in English (Spanish set 1) also improved slightly (from 0% to 27%). No changes were observed in Spanish set 2. Treatment then began with this set as the target words.

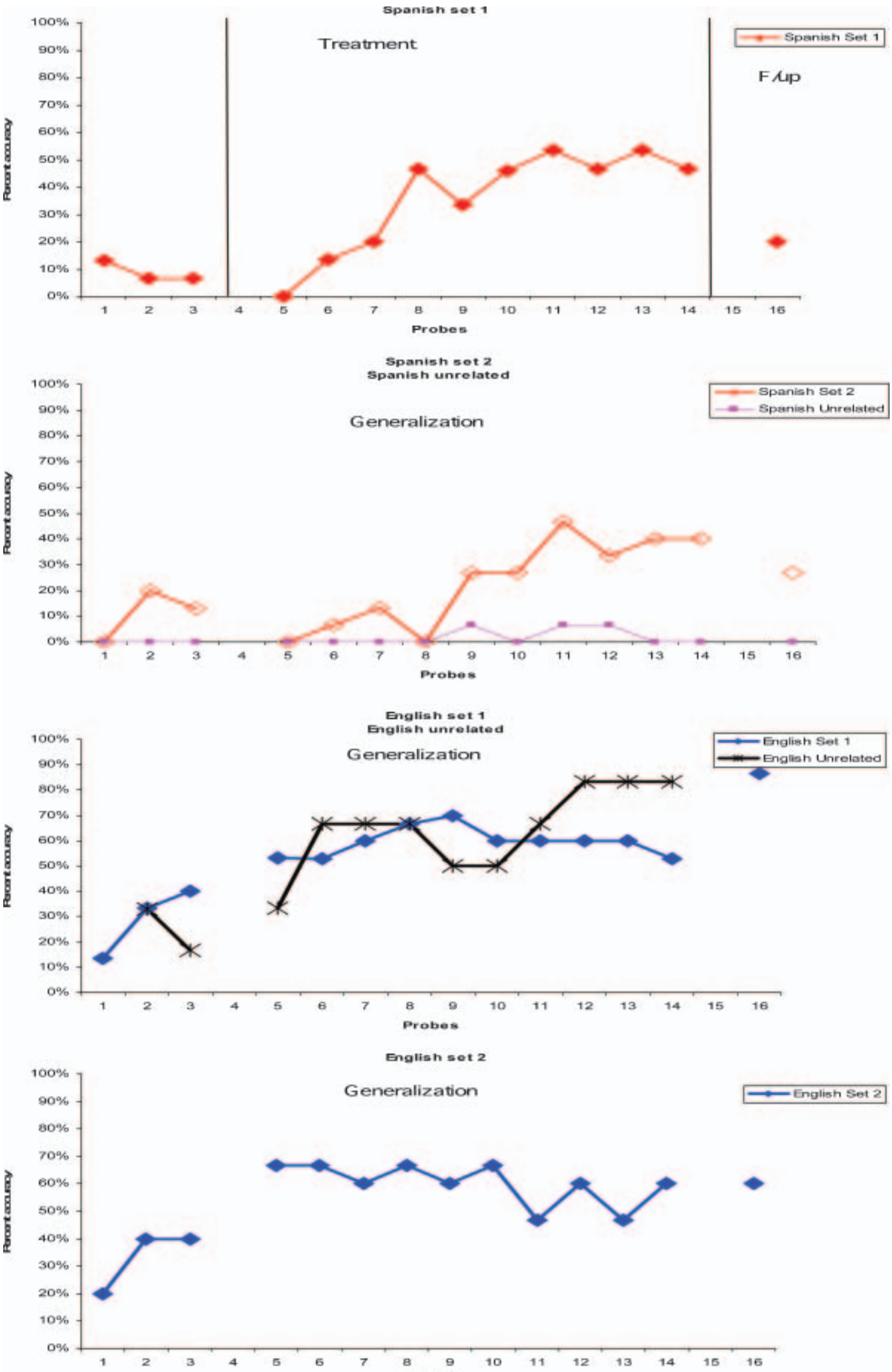


**Figure 1.** Naming accuracy for Participant 1 on Spanish set 1 (trained) and Spanish set 2 (semantically related to set 1), English set 1 (translations of Spanish set 1) and English set 2 (semantically related translations). Control (unrelated) items in English and Spanish are illustrated in the same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

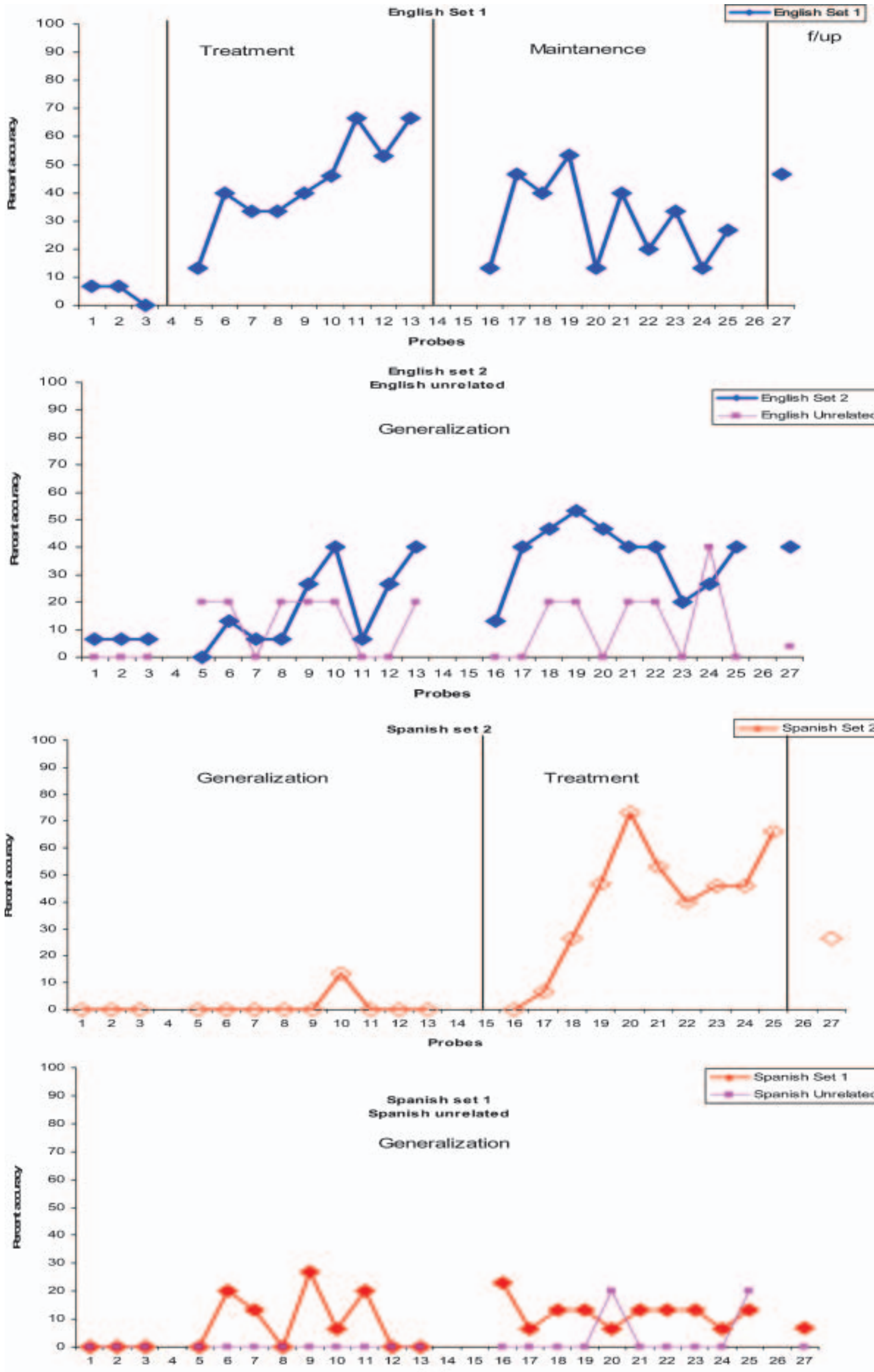


**Figure 2.** Naming accuracy for Participant 2 on English set 1 (trained first) and English set 2 (semantically related to set 1), and Spanish set 2 (trained second) an Spanish set 1 (translations of English set 1). Control (unrelated) items in English and Spanish are illustrated in the same graph as generalisation items.

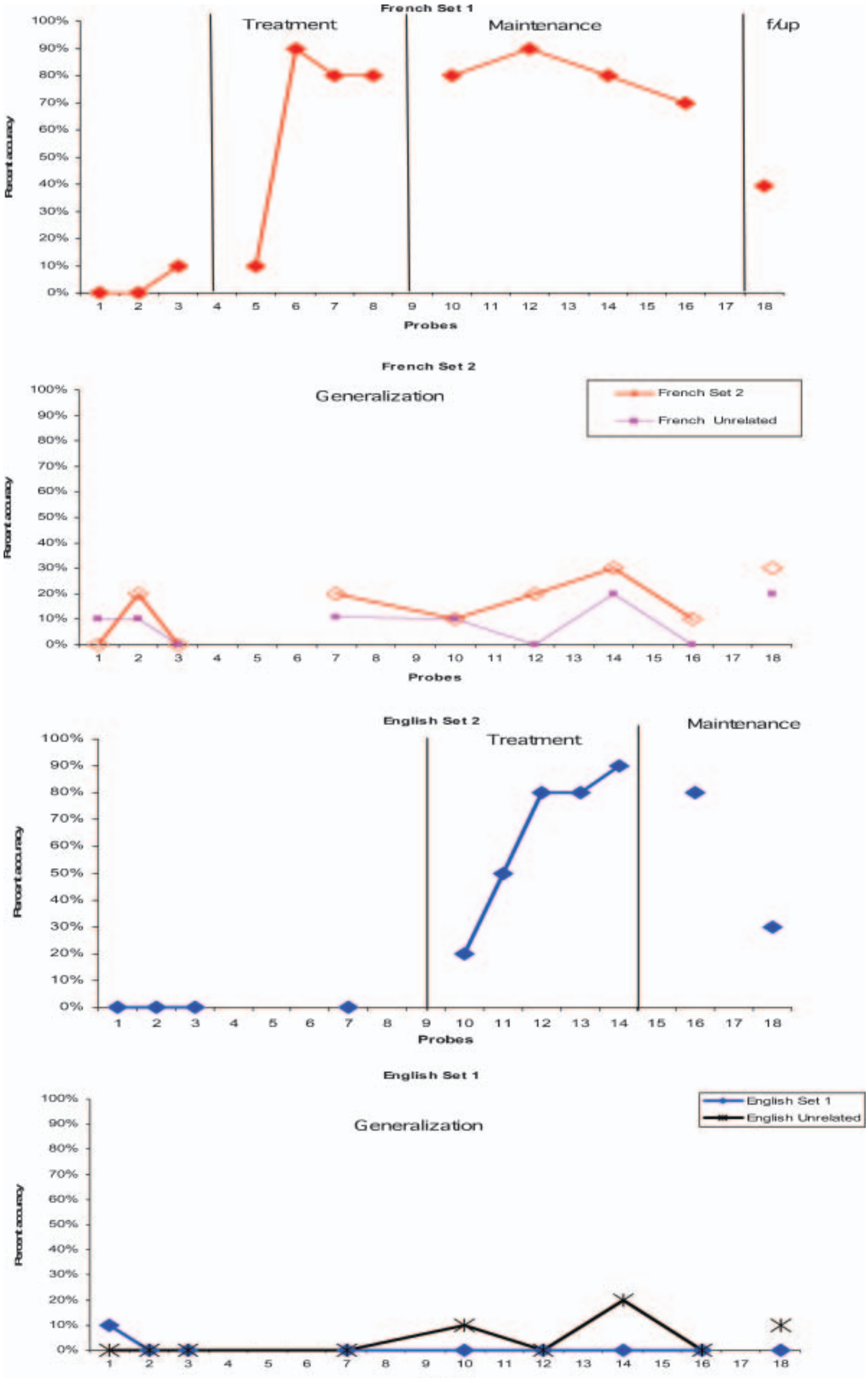


**Figure 3.** Naming accuracy for Participant 3 on French set 1 (trained first) and French set 2 (semantically related to set 1), and English set 2 (trained second) an English set 1 (translations of English set 1). Control (unrelated) items in English and Spanish are illustrated in the same graph as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

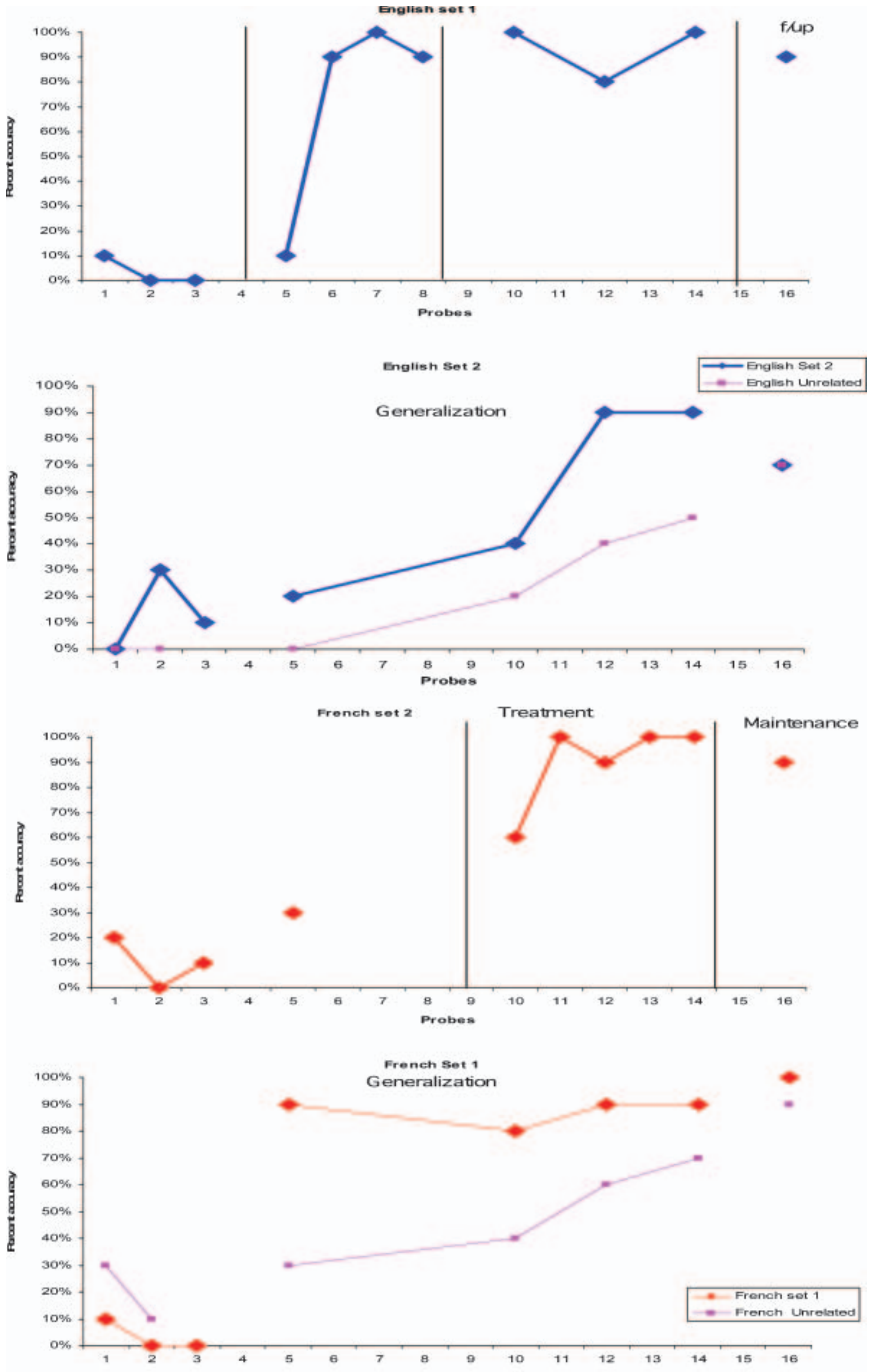


**Figure 4.** Naming accuracy for Participant 4 on English set 1 (trained first) and English set 2 (semantically related to set 1), and French set 2 (trained second) an French set 1 (translations of English set 1). Control (unrelated) items in English and Spanish are illustrated in the same graph as generalisation items.

Performance improved to a high of 73% but was unstable (ES = 8.8). Surprisingly, performance deteriorated in the semantically related but untrained Spanish words (Spanish set 1) but was overall higher than baseline levels (ES = 2.8). Performance on the untrained English translations of the trained items (i.e., English set 2) showed a variable but ultimately higher average performance (36.6%) than in the previous phase (18.5%) or during baseline (ES = 5.5). Also, performance on the previous trained English set 1 declined slightly (ES = −0.8). Scores for unrelated control items in English (ES = 3.2) and Spanish (ES = 0.5) were unchanged.

### *Participant 3*

P3 was initially trained in French (her stronger language) and then in English. After just one session, naming on the trained items (French set 1) improved dramatically (ES = 13.9). No changes were observed on the untrained but related French set 2 or on the English sets. When treatment shifted to English Set 2, performance on these items jumped to 90% (ES = 11.5), with no change in the semantically related untrained English set 1 (ES = −0.57). Performance on the previously trained French set 1 was maintained after therapy at 70% accuracy for those items that were at 80% when treatment was discontinued. Performance on the untrained French set 2 (ES = 0.98) and the unrelated French control items (ES= 0.26) changed little. Lastly, performance on the unrelated English controls items did not change (ES = 1.03). At follow-up approximately 5 weeks after completing treatment, performance had declined to at or near baseline levels. The patient did not attend two scheduled follow-up appointments to check maintenance/ progress at 3 months post-therapy.

### *Participant 4*

P4 was initially treated in English and then in French. Training on English set 1 resulted in a rapid improvement to a high of 100% (ES = 13.5). Performance at that time on the English semantically related words (set 2) did not change nor did performance on French set 2. However, there was improvement in the French translations of English set 1 (i.e., French set 1) to a high of 90% after treatment in English. Performance on the unrelated English and French items did not change after English treatment.

Treatment was then shifted to French set 2 which improved to a high of 100% in four sessions (ES = 5.8). Performance on the untrained, semantically related French set 1 continued to be highly accurate (ES = 14.5). Her score on the untrained English translations of the trained set (English set 2) improved from 40% to 90% (ES = 3.05). Scores on the unrelated French and English sets improved to a high of 70% (ES = 4.0) and 50% (ES = 2.12) respectively. At follow-up 3 months post-discharge, scores on all sets of words were similar to those at the end of therapy.

## Error analysis

For all participants, responses produced during the first baseline session and at the end of treatment for each language (except P1 who only received treatment in Spanish), were coded and are shown in Table 6. P1 showed primarily phonemic and semantic errors in English before Spanish treatment. After treatment in Spanish, the semantic errors remained but there were fewer "no-response" errors. For Spanish stimuli, the main error types were no-responses, neologisms, and semantic errors. After treatment, the main error types were still neologisms, unrelated words, and a few semantic errors.

P2, treated in English then Spanish, showed a variety of errors on English stimuli including no-responses, semantic errors, unrelated words, and neologisms prior to treatment. Even after treatment these error types persisted for the English stimuli. For Spanish stimuli, errors were predominantly no-responses, neologisms, unrelated words, and phonemic errors that were produced before and after treatment. Interestingly, this patient demonstrated more semantic errors in English and more phonemic errors in Spanish. She also showed a tendency to produce semantic errors or the translation of the target in the non-target language.

P3, who was treated first in French then in English, showed predominantly no-responses and semantic errors in English before and after treatment. The pattern was similar for French stimuli: errors were either no-responses or semantic errors prior to treatment. Following treatment, however, semantic errors continued and a few cross-language translations were also observed.

Finally P4, who received treatment first in English and then in French, showed predominantly no responses and semantic errors prior to treatment for both English and French stimuli. Since this patient showed within- and cross-language generalisation, few errors remained post-treatment. These were mainly semantic errors. Interestingly, P3 and P4 made more semantic errors for French stimuli than English stimuli.



## Results on standardised tests administered in both languages

Several of the standardised measures that were administered pre-treatment were administered again post-treatment (see Tables 2 and 3). Scores on the *Boston Naming Test* deteriorated in English but improved in Spanish for P1, whereas scores improved in English and Spanish for P2. Slight fluctuations with no specific trends were observed for P3 and P4 (see Figure 5). Performance on the *BAT* fluctuated for the four participants with no apparent pattern.

# DISCUSSION

The goals of the present study were to test predictions about bilingual language treatment, and to replicate and extend previous work on cross-language generalisation in bilingual aphasia to new patients and a new language combination (English–French). The experimental design and methodology were very similar to the Edmonds and Kiran (2006) model. Results of the present study revealed somewhat different patterns of results for the four patients. After reviewing the results in light of the five hypotheses proposed in the introduction, we conclude with a discussion of the challenges in conducting and interpreting studies of bilingual aphasia treatment.

Each patient responded differently to the SFA treatment in terms of their improvement on the target items, the within- and between-language generalisation and their maintenance. Table 5 summarises the results and some of the relevant linguistic and other variables for the four patients. For each hypothesis, the results break down as follows:

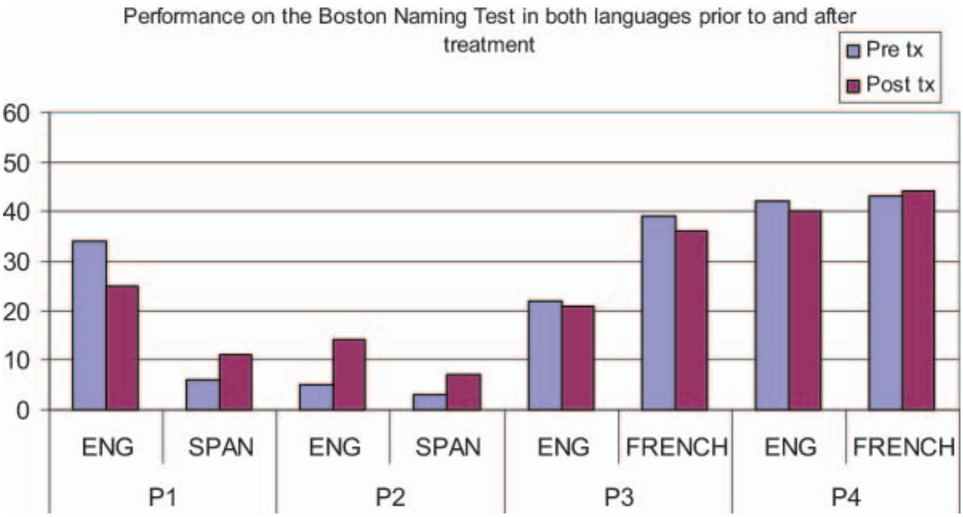


**Figure 5.** Performance on the Boston Naming Test prior to treatment and after treatment for Participants 1–4. For Participants 1 and 2, performance is shown in English and Spanish and for Participants 3 and 4, performance is shown for English and French.

First, gains on trained items in the trained language occurred for all patients, confirming Hypothesis 1 and providing some support for the efficacy of Semantic Feature Analysis. For P3 and P4, the very rapid improvement and the fact that the trained items improved so much more than the untrained ones strongly suggest that the therapy caused the gains. The improvement may stem from hearing/reading the picture names at least as much as from the strengthened semantic networks in these two patients. However, P1 and P2 failed to reach more than 70% accuracy on the trained words despite the relatively large number of treatment sessions and the fact that they were only 6 months and 11 months post-onset. At least for P1, the lack of robust improvements on the trained items may be influenced by the fact that treatment was provided in her (pre-morbidly) weaker and (post-stroke) more impaired language. Also, both P1 and P2 showed more severe language impairments than P3 and P4.

Second, within-language generalisation occurred for P1, P2, and P4 on items semantically related to the words practiced in therapy, as predicted by Hypothesis 2. However, for P3 this generalisation did not occur.

The cross-language generalisation as predicted by Hypothesis 3 occurred, with clear improvements on translations of the target items for only one of the four patients (P4). Even though P1 showed improvements in English subsequent to treatment in Spanish, the broad improvements of P1 on English set 1 and the increasing scores during the baseline phase make it difficult to interpret her results in terms of this hypothesis. For P2, slight improvements were observed for English set 2 when Spanish set 2 was trained, but again, these results were not robust.

Cross-language generalisation for semantically related targets was predicted by Hypothesis 4, and occurred for P4 and to a lesser extent for P1. While there were improvements in English set 2 for P1, high scores during baseline and the improvement in semantically unrelated words weaken any conclusions drawn from this hypothesis.

TABLE 5
Summary of important pre-stroke language history, post stroke language treatment, and patterns of generalisation observed in the four participants

| *Participants* | *P1* | *P2* | *P3* | *P4* |
|---|---|---|---|---|
| MPO | 11 | 6 | 33 | 15 |
| Stronger language (self-report) | English | approx equal | French | approx equal |
| Language of previous tx | English | English | both, more English | almost all English |
| Language of tx this study (Lt) | Spanish | English, then Spanish | French, then English | English, then French |
| Number of tx sessions | 20 in Spanish | 18 in English<br>20 in Spanish | 4 in each language | 4 in each language |
| Reached criterion of ⩾80% | No for tx words. | No for all lists in both languages | Yes, in both tx lists | Yes, in both tx lists |
| Generalisation: within-language | Some, to semantically-related list only | Some during English tx to sem related. Also some, but less, to unrelated words | None | For Eng tx words Eng set 1 to set 2: yes strong. |
| | Improvement on unrelated words | | | Some but less to unrelated words during each tx phase |
| Generalisation: cross-languages | Improvement in English set 2 but questionable | Slight improvement from Spanish set 2 to English set 2 | Slight improvement from English set 2 to French set 2 | Strong from Eng set 1 (tx) to French set 1; slight to unrelated |

MPO: months post onset, tx: treatment, sem: semantic.

TABLE 6
Evolution of errors reported in raw numbers for each patient

| | *BL* | *En Tx* | *Sp Tx* | *BL* | *En Tx* | *Sp Tx* | *BL* | *En Tx* | *Sp Tx* | *BL* | *En Tx* | *Sp Tx* | *BL* | *En Tx* | *Sp Tx* | *BL* | *En Tx* | *Sp Tx* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *English 1* | | | *English 2* | | | *English-UNR* | | | *Spanish 1* | | | *Spanish 2* | | | *Spanish-UNR* | | |
| *P1* | (15) | — | (15) | (15) | — | (15) | (6) | — | (6) | (15) | — | (15) | (15) | — | (15) | (6) | — | (6) |
| No-Response | 1 | — | 2 | 2 | — | 1 | 3 | — | 0 | 5 | — | 1 | 3 | — | 2 | 3 | — | 2 |
| Neologism/Perseveration | 0 | — | 0 | 1 | — | 0 | 0 | — | 0 | 2 | — | 5 | 0 | — | 0 | 1 | — | 3 |
| Unrelated | 0 | — | 1 | 1 | — | 1 | 1 | — | 0 | 0 | — | 1 | 1 | — | 3 | 0 | — | 1 |
| Phonemic Error-TL | 4 | — | 0 | 1 | — | 1 | 0 | — | 0 | 0 | — | 0 | 1 | — | 2 | 0 | — | 0 |
| Semantic Error-TL | 6 | — | 4 | 6 | — | 2 | 1 | — | 1 | 3 | — | 1 | 1 | — | 2 | 1 | — | 0 |
| Circumlocution | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 |
| Mixed | 2 | — | 0 | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 | 1 | — | 0 | 0 | — | 0 |
| Phonemic Error-NTL | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 | 1 | — | 0 | 0 | — | 0 | 0 | — | 0 |
| Semantic Error-NTL | 0 | — | 0 | 0 | — | 1 | 0 | — | 0 | 1 | — | 0 | 2 | — | 0 | 0 | — | 0 |
| Correct-NTL | 0 | — | 0 | 0 | — | 0 | 0 | — | 0 | 2 | — | 0 | 2 | — | 0 | 1 | — | 0 |
| Correct-TL | 2 | — | 8 | 4 | — | 9 | 1 | — | 5 | 1 | — | 7 | 4 | — | 6 | 0 | — | 0 |
| *P2* | (15) | (15) | (15) | (15) | (15) | (15) | (5) | (5) | (5) | (15) | (15) | (15) | (15) | (15) | (15) | (5) | (5) | (5) |
| No-Response | 7 | 0 | 4 | 5 | 3 | 1 | 0 | 0 | 2 | 3 | 0 | 3 | 9 | 5 | 1 | 0 | 0 | 2 |
| Neologism/Perseveration | 1 | 0 | 1 | 2 | 1 | 1 | 3 | 0 | 0 | 7 | 8 | 6 | 2 | 8 | 1 | 2 | 3 | 2 |
| Unrelated | 2 | 1 | 1 | 5 | 4 | 0 | 1 | 1 | 1 | 0 | 2 | 3 | 1 | 1 | 0 | 0 | 0 | 0 |
| Phonemic Error-TL | 1 | 4 | 2 | 1 | 2 | 2 | 0 | 0 | 1 | 3 | 5 | 2 | 1 | 0 | 3 | 0 | 1 | 0 |
| Semantic Error-TL | 3 | 1 | 3 | 1 | 0 | 2 | 1 | 3 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Circumlocution | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Mixed | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Phonemic Error-NTL | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Semantic Error-NTL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| Correct-NTL | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Correct -TL | 1 | 9 | 4 | 1 | 5 | 6 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 10 | 0 | 0 | 1 |

(*Continued*)

TABLE 6
Continued

| | *BL* | *Fr Tx* | *En Tx* | *BL* | *Fr Tx* | *En Tx* | *BL* | *Fr Tx* | *En Tx* | *BL* | *Fr Tx* | *En Tx* | *BL* | *Fr Tx* | *En Tx* | *BL* | *Fr Tx* | *En Tx* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *English 1* | | | *English 2* | | | *English-UNR* | | | *French 1* | | | *French 2* | | | *French-UNR* | | |
| *P3* | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10)* | (10) | (10) | (10) | (10) | (10) | (10) | (9) | (10) |
| No-Response | 10 | 9 | 6 | 9 | 8 | 1 | 8 | 6 | 4 | 6 | 2 | 2 | 6 | 3 | 1 | 5 | 2 | 3 |
| Neologism/Perseveration | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Unrelated | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Phonemic Error-TL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Semantic Error-TL | 0 | 1 | 1 | 1 | 0 | 0 | 2 | 1 | 4 | 3 | 0 | 0 | 3 | 5 | 1 | 2 | 4 | 2 |
| Circumlocution | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 2 | 3 | 2 |
| Mixed | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Phonemic Error-NTL | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Semantic Error-NTL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Correct-NTL | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 |
| Correct-TL | 0 | 0 | 0 | 0 | 1 | 9 | 0 | 1 | 1 | 0 | 8 | 8 | 0 | 2 | 3 | 1 | 0 | 2 |
| *P4* | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) | (10) |
| No-Response | 6 | 0 | 0 | 5 | 0 | 1 | 6 | 0 | 1 | 4 | 0 | 0 | 3 | 0 | 1 | 3 | 0 | 0 |
| Neologism/Perseveration | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Unrelated | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| Phonemic Error-TL | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Semantic Error-TL | 4 | 0 | 0 | 1 | 0 | 2 | 2 | 3 | 4 | 3 | 1 | 1 | 3 | 0 | 4 | 3 | 0 | 4 |
| Circumlocution | 0 | 0 | 0 | 2 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 1 | 1 |
| Mixed | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Phonemic Error-NTL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Semantic Error-NTL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Correct-NTL | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Correct-TL | 0 | 10 | 10 | 0 | 10 | 4 | 0 | 5 | 2 | 1 | 9 | 9 | 2 | 10 | 4 | 3 | 7 | 4 |

Numbers in paranthesis indicate the total number of responses for each participant. Please see text for description of error types. TL: Target Language, NTL: Non Target Language; UNR: Unrelated words, BL: baseline, SP Tx: probes upon completion of treatment in Spanish, En Tx: probes upon completion of treatment in English, Fr: probes upon completion of treatment in English. * Note that one of the responses in this set was not classifiable.

Finally, as predicted by Hypothesis 5, there was no change in scores for the semantically unrelated words in either language for two patients (P2 and P3). However, for P4 and to a certain extent for P1, the "control words" improved, suggesting that for these patients semantic ties between words were not a critical factor in facilitating improved naming.

Results from the error analysis provide further data regarding each patient's response to treatment. For instance, P4 showed the clearest treatment gains and cross-language generalisation patterns and a predominance of semantic errors. In contrast, P2 never consistently responded to treatment in terms of acquisition and generalisation and showed a variety of errors including neologisms, unrelated words, and phonemic errors that did not change as a function of treatment. P3 showed an increase in correct responses in the non-target language and semantic errors even though she did not demonstrate generalisation. Interestingly, this patient was French dominant pre-morbidly and showed an increase in French responses after English treatment (Correct NTL); but an increase in English responses subsequent to English treatment. Lastly, P1 (English-dominant) showed a decrease in the number of semantic errors in English but not in Spanish even though she received treatment in Spanish.

Both the errors and the generalisation trends are consistent with models referred to in the introduction that posit lexical-semantic connections between L1 and L2 and activation of lexical items in both languages, during work in one language. Specifically, the semantic-based naming treatment improved naming of trained items and semantically related items in the trained language and, in P4 only, to translations in the untrained language. In those cases when naming was incomplete (e.g., P4) errors were predominantly semantic errors or translations in the non-target language. In future studies it would be interesting to see if this error pattern is associated with stronger/easier generalisation. These results, however, are not conclusive in terms of the locus of transfer between the two languages in each patient, so it is impossible to know whether treatment strengthened the direct connections between the translation equivalents in the two languages, and/or indirect connections via the semantic level. This issue should be the focus of future bilingual treatment studies.

Given these limitations, the theoretical conclusions that can be drawn from the present data are tentative. If all four patients had shown similar patterns of results or if the extent and direction of generalisation had varied clearly as a function of the patients' language histories, stronger conclusions would have been possible. However, the current data can only be said to indirectly support the basic tenets of theoretical models proposed by de Groot (1992) and Kroll and Stewart (1994) in as much as they do not contradict them. An important and very positive finding from a clinical standpoint is that the connections are manipulable as a function of treatment.

From a clinical standpoint, these results provide several insights into bilingual aphasia rehabilitation. First, they remind us of how variable the performance of patients can be. As is the case for monolingual patients, some patients respond well to treatment, while other apparently similar patients show little or no improvement (e.g., Laganaro, Di Pietro, & Schnider 2006; Law, Wong, Sung, & Hon, 2006; Marini, Caltagirone, Pasqualetti, & Carlomagno, 2007; Nettleton & Lesser, 1991; Wierenga et al., 2006). After many sessions, P1 and P2 still did not reach 80% correct (12/15), whereas P3 and P4 both jumped to 80% (8/10) or more on their second or third therapy session. Perhaps the relatively mild aphasia of P3 and P4 explains this, and perhaps P2 failed to reach criterion because her aphasia was relatively severe. Also, the larger number of target items for P1 and P2 compared to P3 and P4 may have made it harder to achieve the goal of 80%, although the larger number gives a more representative sample of naming behaviour.

Second, the results remind us that we still do not understand why the performance of only some patients generalises to untrained items. One possible interpretation of the present results is that they replicate and extend those of Edmonds and Kiran (2006). The results for P1 and P4 could even be interpreted as supporting the tentative explanation for the 2006 results; i.e., that treatment in a patient's weaker language is more likely to facilitate generalisation to untrained stimuli than treatment in the patient's stronger language. However, there are other plausible interpretations that cannot be ruled out. The various results for these four patients could be interpreted in terms of their bilingualism (strong/weak languages; age of acquisition of each language; patterns of use). On the other hand, the literature on aphasia treatment in monolingual patients also shows that some patients generalise gains much more than others. So the failure of P3 to generalise from her stronger language to her weaker may have nothing to do with her bilingualism or the stimuli used. Similarly, the generalisation shown by P1 (at 6 months post-onset) may be partly due to spontaneous recovery. Cognitive factors such as executive function have been suggested as a possible explanation for the differing effects of treatment across patients (e.g., Fillingham, Sage, & Lambon Ralph, 2005; Law, Yeung, & Chiu, 2008). Future studies of bilingual patients might be easier to interpret if they include measures of problem solving and other cognitive functions. One way to encourage generalisation is to provide a large number of training sessions (e.g., Raymer, Kohen, & Saffell, 2006) and a larger number of stimuli. Yet after only four sessions in each language, and using only 10 target words, P4 showed strong generalisation within and between languages.



## Methodological issues to consider in future studies

The results from the present study raise several issues that cannot be conclusively resolved with the current data and need to be addressed systematically in future studies. For instance, self-reports of proficiency reflect each patient's subjective judgements about their abilities. Although in this study no obvious inconsistencies were found between their language histories, reported age of acquisition, and degree of accentedness in L2, more studies about the validity of self-reports are required. For example, two recent studies showed that participants' self-rated proficiency did not reflect their performance on overt naming (Sebastian & Kiran, 2007) and during semantic priming (Kiran & Lebel, 2007). On the Boston Naming Test, scores correlated modestly with self-rated proficiency in verbal expression at the group level but large variability within each group was found, despite the uniformly high self-ratings of proficiency (Roberts, Garcia, Desrochers, & Hernandez, 2002). There is currently more support for the validity of self-ratings at the group level than for their validity in predicting performance for individuals (see Roberts, 2008, for a review).

A related issue concerns self-reports of age of acquisition for each language. It is difficult to estimate the age of first exposure in older adults. Some do not remember exactly what year English/French/Spanish classes began in school and, even when asked, are unsure about passive exposures through TV, radio, stories; this makes it difficult to reliably classify them and to draw up sets of stimuli that might reflect age of acquisition. It is not clear how important this measure may be for naming recovery, but there is some work suggesting that, especially for priming tasks, and tasks that use reaction time as the dependent variable, there may be word-by-word effects for age of acquisition (e.g., Barry, Johnston, & Wood, 2006; Izura & Ellis, 2002).

One of the difficult aspects of this study was selecting equivalent stimuli in each language. The normal procedure of balancing lists for word length, published word frequency, and for typical age of acquisition is complicated for bilingual speakers. Given Grosjean's (1998) complementarity principle, it is unclear to what extent published word frequency lists for any given language accurately reflect the word frequency or word difficulty for bilingual speakers (Roberts, 2008; Roberts et al., 2002). Even for monolingual speakers, frequency databases that have not separately counted each meaning of homonyms such as *bit* and *mouth* may give very misleading frequency values. Finally, the different samples (written language, Internet, TV shows, newspapers, etc.) used in the word frequency databases and the different ways of calculating frequencies mean that comparing frequencies of a given word in two databases may be invalid.

The validity and reliability of semantic relatedness scores also warrant further study. In the rating exercise, the test–retest reliability was tested by including some pairs more than once, often in reverse order. Mean ratings varied by up to 1 point on a 5-point scale. The extent to which ratings vary with the age and with the level of bilingualism of the raters is unknown. In future studies it would be worthwhile to deliberately vary the degree of relatedness to see if within and/or between language generalisation correlates with rated strength of the semantic ties.

We recognise that the WAB and BNT are not designed to measure aphasia in bilingual speakers. Further, aspects of the BAT require further psychometric scrutiny (e.g., Muñoz & Marquardt, 2008). The ultimate goal of treatment is to effect changes in both languages of bilingual individuals that extend beyond just improvements in naming. Until standardised assessment tools are validated across languages, our ability to make any meaningful interpretations of scores on these tests regarding the severity of aphasia in each language and to measure changes over time in each language is limited.

## Conclusion

The present study is the first replication of a bilingual aphasia treatment study across two languages. There are several strengths in the methodology of the study. First, all four participants were relatively similar in age and were all females. The detailed language histories for all patients included comparisons of language use and language history patterns across two pairs of languages. Further, multiple baselines were obtained for all patients and the order of languages treated was counter-balanced across three participants. Consequently the results of the present study allow us to draw some tentative theoretical and clinical conclusions. The present results demonstrate the feasibility of using semantic treatment to facilitate lexical retrieval and generalisation to semantically related untrained items across three different languages. Treatment efficacy was not uniform for all patients, but all patients improved their naming ability. The pattern of improvement in three of the four patients is strongly linked to the phases and targets of the treatment. These results are somewhat equivocal about the extent of transfer to untrained languages, but the treatment has the potential to benefit items in the untrained language. As better methods are developed for selecting stimuli, controlling for more factors related to those stimuli and for grouping patients based on their linguistic histories, a clearer picture will emerge.

## REFERENCES

Barry, C., Johnston, R. A., & Wood, R. F. (2006). Effects of age of acquisition, age, and repetition priming on object naming. *Visual Cognition*, *13*, 911–927.

Baudot, J. (1992). *Fréquence d'utilisation des mots en français écrit contemporain*. Montréal: Les Presses de l'Université de Montréal.

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: lessons learned from the aphasia literature. *Neuropsychological Review*, *16*, 161–169.

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

Busk, P. L., & Serlin, R. (1992). Meta analysis for single case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single case research design and analysis: New directions for psychology and education*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Coelho, C. A., McHugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, *14*, 133–142.

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 and Cognition*, *2*, 231–244.

Costa, A., La Heij, W., & Navarrete, E. (2006). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, *9*, 137–151.

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

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197.

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., & Kiran, S. (2006). Effect of semantic based treatment on cross linguistic generalisation in bilingual aphasia. *Journal of Speech, Language and Hearing Research*, *49*, 729–748.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005). Further explorations and an overview of errorless and errorful therapy for aphasic word-finding difficulties: The number of naming attempts during therapy affects outcome. *Aphasiology*, *19*, 597–614.

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

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

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*, 131–149.

Haarbauer-Krupa, J., Moser, L., Smith, G., Sullivan, D., & Szekeres, S. F. (1985). Cognitive rehabilitation therapy: Middle stages of recovery. In M. Ylvisaker (Ed.), *Head injury rehabilitation: Children and adolescents* (pp. 287–310). San Diego, CA: College Hill Press.

Izura, C., & Ellis, C. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicólogica*, *23*, 245–281.

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

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

Kay, J., Lesser, R., & Coltheart, M. (1992). *The Psycholinguistic Assessment of Language Processing in Aphasia (PALPA)*. Hove, UK: Lawrence Erlbaum Associates Ltd.

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

Kiran, S., & Lebel, K. (2007). Crosslinguistic semantic and translation priming in normal bilingual individuals and bilingual aphasia. *Clinical Linguistics and Phonetics*, *21*, 277.

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*, 773–787.

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.

Laganaro, M., Di Pietro, M., & Schnider, A. (2006). What does recovery from anomia tell us about the underlying impairment: The case of similar anomic patterns and different recovery. *Neuropsychologia*, *44*, 534–545.

Law, S-P., Wong, W., Sung, F., & Hon, J. (2006). A study of semantic treatment of three Chinese anomic patients. *Neuropsychological Rehabilitation*, *16*, 601–629.

Law, S-P., Yeung, O., & Chiu, K. M. Y. (2008). Treatment for anomia in Chinese using an ortho-phonological cueing method. *Aphasiology*, *22*, 139–163.

Marini, A., Caltagirone, C., Pasqualetti, P., & Carlomagno, S. (2007). Patterns of language improvement in adults with non-chronic, non-fluent aphasia after specific therapies. *Aphasiology*, *21*, 164–186.

Muñoz, M. L., & Marquardt, T. P. (2008). The performance of neurologically normal bilingual speakers of Spanish and English on the short version of the Bilingual Aphasia Test. *Aphasiology*, *22*, 3–19.

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

Nettleton, J., & Lesser, R. (1991). Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model. *Journal of Neurolinguistics*, *6*, 139–157.

Paradis, M. (1989). *Bilingual Aphasia Test*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

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

Raymer, A. M., Kohen, F. P., & Saffell, D. (2006). Computerised training for impairments of word comprehension and retrieval in aphasia. *Aphasiology*, *20*, 257–268.

Roberts, P. M. (1998). Clinical research needs and issues in bilingual aphasia. *Aphasiology*, *12*, 119–130.

Roberts, P. M. (2008). Issues in assessment and treatment in bilingual and multicultural populations. In R. Chapey (Ed.). *Language intervention strategies in adult aphasia* (5th ed., pp. 245–276). Baltimore: Lippincott, Williams & Wilkins.

Roberts, P. M., & Deslauriers, L. (1999). Picture naming of cognate and non-cognate nouns in bilingual aphasia. *Journal of Communication Disorders*, *32*, 1–23.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, *16*, 635–645.

Roberts, P. M., & Kiran, S. (2007). Assessment and treatment of bilingual aphasia and bilingual anomia. In A. Ardila & E. Ramos (Eds.), *Speech and language disorders in bilinguals* (pp. 109–131). New York: Nova Science Publishers.

Roberts, P. M., & Shenker, R. C. (2007). Assessment and treatment of stuttering in bilingual speakers. In R. F. Curlee & E. G. Conture (Eds.), *Stuttering and related disorders of fluency* (3rd ed., pp. 183–210). New York: Thieme Medical Publishers.

Sebastian, R., & Kiran, S. (2007, November). *Semantic processing in Hindi–English bilinguals using fMRI*. Poster, ASHA Convention, Boston, MA.

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, *51*, 381–398.

Wierenga, C. E., Maher, L. M., Moore, A. B., White, K. D., McGregor, K., & Soltysik, D., et al. (2006). Neural substrates of syntactic mapping treatment: An fMRI study of two cases. *Journal of the International Neuropsychological Association*, *12*, 132–146.

# APPENDIX

Stimuli and average frequency values (standard deviations in parenthesis) for each patient

| *English set 1* | *Spanish set 1* | *English set 2* | *Spanish set 2* | *Average semantic relatedness* [1] | *English UR* | *Spanish UR* |
|---|---|---|---|---|---|---|
| *Participant 1* | | | | | | |
| Ant | Hormiga | Spider | Arana | 1.8 | Blanket | Manta |
| Razor | Rastrillo | Soap | Jabón | 2.7 | Wallet | Cartera |
| Leg | Pierna | Arm | Brazo | 1.6 | Pitcher | Jarra |
| Cane | Bastón | Umbrella | Paraguas | 3.0 | Sword | Espada |
| Sheep | Borrego/ Oveja | Deer | Venado/Ciervo | 2.1 | Rug | Alfombra |
| Cloud | Nube | Lightning | Relámpago | 2.3 | | |
| Eagle | Aguila | Owl | Buho | 2.0 | | |
| Raccoon | Mapache | Skunk | Zorrillo | 2.0 | | |
| Shark | Ttiburon | Whale | Ballena | 1.7 | | |
| Snail | Caracol/ Coconito | Worm | Gusano | 2.0 | | |
| Nun | Monja | Teacher | Maestra | 2.0 | | |
| Shelf | Estante | Hook | Gancho | *3.0* | | |
| Stool | Taburete | Counter | Mostrador/ Ventanilla | 2.3 | | |
| Wheelbarrow | Carretilla | Dustpan | Recogedor | 2.3 | | |
| Mouth | Boca | Mus- tache | Bigote | 2.3 | | |
| *Participant 2* | | | | | | |
| Cabbage | Repollo | Radish | Rabano | 1.8 | Mustache | Tenedor |
| Dustpan | Recogedor | Vacuum cleaner | Aspiradora | 2.0 | Garlic | Buho |
| Forehead | Frente | Chin | Menton | 3.0 | Deer | Lambriz |
| Raccoon | Mapache | Skunk | Zorillo | 2.0 | Umbrella | Naranja |
| Razor | Rasador | Soap | Jabon | 2.7 | Wallet | Meastra |
| Necklace | Collar | Ring | Anillo | | | |
| Shark | Tiburon | Whale | Ballena | 1.7 | | |
| Counter | Mostrador | Hook | Gancho | 2.3 | | |
| Stool | Taburette | Shelf | Estante | 2.3 | | |
| Wheelbarrow | Carretilla | Rake | Rastrillo | 2.7 | | |
| Newspaper | Periodico | Magazine | *Revista* | 1.9 | | |
| Robe | Bata | Coat | Abrigo | 2.3 | | |
| Wrench | Perica | Drill | Taladro | 1.6 | | |
| Lightning | Relampago | Cloud | *Nube* | *2.3* | | |
| Spider | Arana | Ant | Hormiga | 1.8 | | |

| *English set 1* | *French set 1* | *English set 2* | *French set 2* | *Average semantic relatedness*[2] | *English UR* | *French UR* |
|---|---|---|---|---|---|---|
| *Participant 3* | | | | | | |
| Button | Macaron | Badge | Écusson | 1.5 | Landing | Palier |
| Ditch | Fossé | Puddle | Flaque | 2.3 | Cheetah | Guépard |
| Dragonfly | Libellule | Spider | Araignée | 2.3 | Clasp | Fermoir |
| Greenhouse | Serre | Factory | Usine | 1.9 | Propellers | Hélices |
| Hinge | Charnière | Bolt | Boulon | 2.3 | Runway | Piste d'atterisage |
| Squash | Courge | Eggplant | Aubergine | 1.2 | Pushpin | Punaise |
| Stirrups | Étriers | Bit | Mors | 1.8 | Speaker | Haut parleur |
| Stopwatch | Chronomètre | Sundial | Cadran solaire | 2.5 | Helmet | Casque |
| Swan | Cygne | Partridge | Perdrix | 2.1 | Wind turbines | Éoliennes |
| Walrus | Morse | Whale | Baleine | 1.6 | Spokes | Rayons |
| *Participant 4* | | | | | | |
| Beetle | Coccinelle | Caterpillar | Chenille | 2.4 | Carnation | Oeillet |
| Cuff | Poignet | Tights | Collants | 3.0 | Firehall | Caserne |
| Danish | Danoise | Waffle | Gauffre | 2.0 | Partridge | Perdrix |
| Fin | Nageoire | Flipper | Palme | 1.5 | Speaker | Haut parleur |
| Bit | Mors | Blinders | Oeillères | 2.3 | Spokes | Rayons |
| Bolt | Boulon | Hinge | Charnière | 2.1 | Step ladder | Escabeau |
| Paperclip | Trombone | Staples | Agraphes | 1.4 | Clasp | Fermoir |
| Stopwatch | Chronomètre | Sundial | Cadran solaire | 2.4 | Landing | Palier |
| Shingles | Bardeaux | Shutters | Volets | 2.5 | Starfish | Étoile de mer |
| Briefcase | Porte-documents | Glasses case | Étui à lunettes | 2.2 | Button | Macaron |

[1]4 point scale; 1 = maximum relatedness. [2]5 point scale; 1 = maximum relatedness.