

Methodological Considerations in Code-Switching Research

Kay González-Vilbazo, Laura Bartlett, Sarah Downey, Shane Ebert, Jeanne Heil,
Bradley Hoot*, Bryan Koronkiewicz & Sergio Ramos
University of Illinois at Chicago & DePaul University*

Abstract

This article addresses methodological concerns in research on grammatical aspects of code-switching. Data from code-switching have the potential for a unique contribution to linguistics by giving us access to combinations of linguistic features that may be difficult (or impossible) to observe in monolingual data. Nonetheless, the use of code-switching data for linguistic inquiry is not without issues. In this paper, we focus on three methodological questions specific to code-switching research: (i) project design, (ii) experimental procedure and (iii) participant selection. Drawing on experimental data from both published works and in-progress projects, we highlight potential solutions to each methodological challenge, concluding that several solutions are often required to mitigate the impact of confounding variables. In line with previous work (e.g. Grosjean 1998, Gullberg, Indefrey & Muysken 2009), we suggest that researchers clearly report on their methodology. Our overall goal is to contribute to a dialogue on best practices in code-switching research.

1. Introduction

In this article we explore methodological questions in studying grammatical aspects of code-switching (CS), i.e. the simultaneous use of two languages within a discourse by bilingual speakers. Here, we are primarily interested in CS as an expression of a bilingual speaker's I-language. I-language is defined by Chomsky (1986) as the mentally represented linguistic knowledge of a native speaker. This knowledge is reflected in the competence of every native speaker. One of the goals of linguistics is to understand the properties of a speaker's competence in order to have access to the fundamental principles of the human language faculty. To this effect, linguists build language models that can generate all and only those sentences considered acceptable by native speakers of a language. While historically these models have been based on data gathered from monolingual speakers, bilingual speakers also have grammatical competence, i.e. they retain clear intuitions about the acceptability or unacceptability of code-switched

utterances (Toribio 2001a, b). This competence has been considered suggestive of largely autonomous and differentiated (not mixed) systems, especially in bilingual children (Genesee 1989, Meisel 1990 among others). Therefore, CS falls within the range of possible human languages. As a result, language models concerned with describing the human language faculty should also account for, and draw from, CS data.

One way in which CS is a useful source for linguistic data is by giving us access to combinations of linguistic features that may otherwise be difficult (or impossible) to observe in monolingual data (González-Vilbazo & López 2012). Though potentially fruitful, the use of CS data for linguistic inquiry is not without issues. The following example serves to illustrate the advantages of using CS data to test general linguistic theories.

This example is taken from a study on sluicing featuring CS between German and Spanish (González-Vilbazo & Ramos forthcoming). Sluicing, sometimes known as TP-ellipsis, is an example of an elliptical construction, where part of a clause is missing, but the meaning can still be reconstructed. Consider (1).¹

- (1) John threatened someone, but I don't know [_{CP} who_i < [_{TP} John threatened t_i] >].

In (1), the TP in angle brackets is 'sluiced' (not pronounced), and accounting for the fact that its meaning can still be reconstructed has been the subject of significant research (e.g. Merchant 2001, Ross 1969). Two main approaches diverge with respect to the connection between the sluiced site and the antecedent. According to one line of research, the sluiced part of the sentence must be semantically equivalent to the antecedent or at least entailed via *e-givenness* (van Craenenbroeck 2010, Merchant 2001). The second line of research argues that the sluiced constituent is also required to be morpho-syntactically identical to the antecedent (Chung 2006, Merchant 2008).

The question is how we can bring empirical evidence to bear on this theoretical debate. If we could see that a morphosyntactic feature was clearly involved, we could provide evidence against a semantic-only identity condition. Morphological case may be just such a feature. A well-known observation in sluicing (Ross 1969) is that the *wh*-phrase (the remnant) remaining from the non-pronounced clause must bear the case that it otherwise would if the full structure of the clause were audible. Let us look at (2) and (3) for instances of sluicing where morphological case (in bold) is visible.

- (2) Juan hat jemandem gedroht, aber ich weiß nicht [_{CP} **wem** < [_{TP}] >].
 (German)
 J has someone.DAT threatened but I know not **who**.DAT
 'Juan threatened someone, but I don't know who.'

- (3) Juan amenazó a alguien, pero no sé [CP **a quién** < [TP] >].
 (Spanish)
 J threatened ACC someone but not know.1SING ACC **who**
 ‘Juan threatened someone, but I don’t know who.’

The German word *drohen* ‘threaten’ assigns dative case to its complement. This case is realized overtly in German on the wh-remnant *wem*. The Spanish verb *amenazar*, which also means ‘threaten,’ assigns accusative case. Given that the wh-phrase is overtly marked for case in German, we can combine both languages in one code-switched utterance and put both approaches to sluicing to the test. Consider now (4) and (5).²

- (4) Juan amenazó a alguien, *aber ich weiß nicht* [CP **wen** < [TP] >].
 J threatened ACC someone but I know not **who.ACC**
 ‘Juan threatened someone, but I don’t know who.’
- (5) *Juan amenazó a alguien, *aber ich weiß nicht* [CP **wem** < [TP] >].
 J threatened ACC someone but I know not **who.DAT**
 ‘Juan threatened someone, but I don’t know who.’

As the examples above show, even though the remnant wh-phrase is in German and embedded in a German clause, it nevertheless must bear accusative case (4), reminiscent of the Spanish example in (3). In this construction, dative case is not an option for the wh-remnant (5). This suggests that the verb in the sluiced clause must have the property of assigning accusative case, rather than dative case, i.e. the sluiced verb must have the same case-assigning properties as the verb in the antecedent clause. This conclusion provides support for the stronger hypothesis that the antecedent and (at least some part of) the sluiced clause must be morphosyntactically, and not only semantically, identical. This is an example of how CS can provide a kind of empirical argument that monolingual data cannot.

Having seen how CS data can benefit linguistic inquiry, the purpose of this work is to draw attention to some methodological issues concerning the gathering and analysis of CS data, as well as to provide illustrative examples and offer some possible solutions to potential problems. We address three areas in this paper: project design, experimental procedure and participant selection. It is worth noting that this work is by no means an attempt to cover the breadth of issues related to linguistic methodology, nor do we offer ultimate solutions. Rather, the solutions offered here exemplify some of the ways in which our laboratory has dealt with these concerns, and as such these solutions are meant to start a conversation on best practices and foment discussion as to how better to proceed in the study of CS.

Here we focus on the methodology of gathering data via experimental acceptability judgment tasks (sometimes referred to as grammaticality judgment

tasks, or GJTs) in order to build models of I-language using CS data. Thus, within the larger set of linguistic methods, our focus is only on a small subset within experimental methods, as can be seen in Figure 1. For a more comprehensive review of research techniques for the study of code-switching see Gullberg, Indefrey & Muysken (2009).

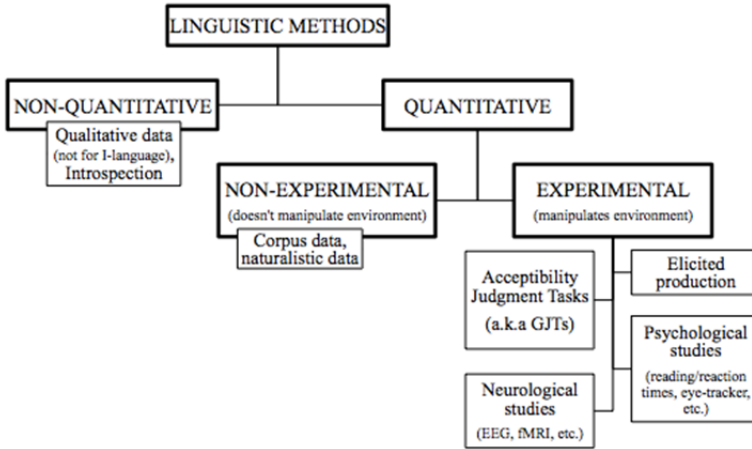


Figure 1. Methodology in linguistics

As shown in Figure 1, we can roughly distinguish between non-quantitative and quantitative methods. Within the latter, we focus on a subset of methods that manipulate the environment,³ i.e. that are experimental in nature, under which we subsume GJTs. While there is a debate about the use and/or validity of GJTs in the field, we will not get into this larger discussion (see, for instance, Schütze 1996 for more details).

The relevance of the present work is based on the fact that there are still relatively few experimental studies of grammatical aspects of CS. While there is a burgeoning research agenda on experimental CS (e.g. Bartlett & González-Vilbazo forthcoming a, Bartlett & González-Vilbazo forthcoming b, González-Vilbazo & Koronkiewicz submitted and González-Vilbazo & Ramos forthcoming), the seminal works in the field (e.g. Belazi, Rubin & Toribio 1994, Poplack 1980 and Zentella 1997) do not use this method. Many studies on CS have relied either on introspection or on naturalistic data such as corpora, both of which have their own methodological advantages and problems.

The presentation in Table 1 explicitly states our position that experimental approaches combine some of the advantages of both introspection and corpus work. The experimental approach that we advocate here combines two positive features of

what we dub *introspection*, by which we mean the common method of the researcher him/herself judging the acceptability of a sentence. First, it gives us access to negative evidence, without which it is not possible to build a realistic picture of a speaker's I-language (Chomsky 1957). Second, an experimental design allows for better control of the stimuli used as well as the environmental conditions under which GJTs take place. Of particular interest for the study of grammatical aspects of CS is the ability to control switch sites, the language pairs involved and the features of the lexical items involved. Finally, the experimental approach also shares an advantage with naturalistic data collected in corpora in that it allows for quantitative data analysis, which is generally not the case with introspection.

	Introspection	Corpus	Experimental
Negative evidence	✓	✗	✓
Control stimuli	✓	✗	✓
Quantitative	✗	✓	✓

Table 1. Methodologies used in code-switching research

Examples of studies using introspection are, among many others, DiSciullo, Muysken & Singh 1986 and Belazi, Rubin & Toribio 1994. Examples of bilingualism-based corpora include the Siarad Welsh-English corpus, the Patagonia Welsh-Spanish corpus, and the Miami Spanish-English corpus, from the Canolfan ESRC Centre. Finally, examples of experimental methodology used in theoretical studies of code-switching are, among others, Bartlett & González-Vilbazo (forthcoming b) and González-Vilbazo & Ramos (forthcoming).

In what follows we will focus on the subset of linguistic methodologies that utilize experimental methods, for the reasons outlined above, and address some issues specific to using experimental GJTs to investigate grammatical aspects of CS. The paper is organized as follows: we begin in Section 2 with project design, then move on to experimental procedure in Section 3 and finally to questions of participant selection in Section 4. Section 5 summarizes our findings and states our conclusions.

2. Project design

2.1 Introduction

Of the many issues regarding project design, we will limit ourselves to considering how to design stimuli effectively for a CS experiment and how to mitigate two

potential confounds: the selection of lexical items and the modality of presentation of stimuli.

2.2 Stimuli

First, let us turn to the issue of designing stimuli to test particular claims in linguistic theory. In a sense, this is the crux of the enterprise and the way in which CS can uniquely contribute to linguistic theory. As was shown in the introduction, by combining two languages that are different in the relevant features of study, we can unveil characteristics of the phenomenon under study that may not be readily available through monolingual data. We can thus test hypotheses that hinge on one feature or on a precise combination of features. This involves both choosing an appropriate language pair and creating stimuli so that the code-switch gives us information about different combinations of features.

To illustrate this point, let us briefly look at an example: a project on wh-movement in Taiwanese-Spanish CS (González-Vilbazo, Bartlett, Ebert & Vergara in prep.). Among the many differences between these two languages, this project crucially hinges on two important differences found in questions: the position of wh-phrases and verb-subject inversion. Examples (6) and (7) illustrate the different positions of wh-constituents in questions for both languages.

- (6) ¿[Cuál de esas bombillas]_i compró *t_i* Mirta? (Spanish)
which of these *mate*⁷ straws bought M
 ‘Which of those *mate* straws did Mirta buy?’
- (7) Mirta be-tio hia-e tue-tsit pun tse? (Taiwanese)
 M bought those **which** CL book
 ‘Which of those books did Mirta buy?’

Whereas Spanish exhibits wh-movement (6), i.e. fronting of the wh-constituent in the question, Taiwanese is a wh-in-situ language,⁸ i.e. the wh-constituent remains in its base position (7). Further, (6) also shows how in Spanish, typically an S-V language, there is inversion between the verb and the subject, something that is not the case in Taiwanese (7). A question for which CS can provide new insight, then, revolves around the very nature of wh-movement. By combining these typologically different languages, we can explore what element(s) might be involved in triggering wh-movement and/or subject-verb inversion. To do so, we create a paradigm with all relevant combinations of the two languages. Table 2 shows a simplified paradigm that only takes four factors into consideration: the language of the verb, the language of the wh-phrase, whether the wh-constituent has been fronted and finally whether there is subject-verb inversion.

	Verb	Wh-Phrase	Movement	Inversion
Mirta <i>be-tio</i> cuál de esas bombillas? M bought which of these mate straws 'Which of these <i>mate</i> straws did Mirta buy?'	<i>TW</i>	SP	N	N
<i>Be-tio</i> Mirta cuál de esas bombillas? Cuál de esas bombillas Mirta <i>be-tio</i> ? Cuál de esas bombillas <i>be-tio</i> Mirta?	<i>TW</i>	SP	N	Y
Mirta compró <i>hia-e tue-chit riab ba-tzang</i> ? M bought those which CL rice dumpling 'Which rice dumpling did Mirta buy?'	SP	<i>TW</i>	N	N
Compró Mirta <i>hia-e tue-chit riab ba-tzang</i> ? <i>Hia-e tue-chit riab ba-tzang</i> Mirta compró? <i>Hia-e tue-chit riab ba-tzang</i> compró Mirta?	SP	<i>TW</i>	N	Y
	SP	<i>TW</i>	Y	N
	SP	<i>TW</i>	Y	Y

Table 2. Paradigm for Taiwanese(*TW*)-Spanish(*SP*) wh-stimuli

By investigating the acceptability judgments provided by the participants for each sentence in the table, it might be possible to better understand what triggers wh-movement. More specifically, these stimuli allow us to isolate lexical items and their associated features and to study what effects the combination of these factors have on the acceptability of the sentence. As Table 2 shows, it is possible to control not only the language of the lexical items that may factor in wh-movement (verb, wh-phrase), but also the typological differences (inversion) that may, directly or indirectly, affect the availability of movement in a language. Depending on the results, it may be necessary to control further factors in the stimuli design. The point is, however, that an experiment of this sort may point us to the features responsible for a linguistic phenomenon, such as wh-movement, in a more nuanced way than monolingual data may.

2.3 Confounds

2.3.1 Lexical items

Although there are advantages to using CS to test linguistic hypotheses, different methodological issues may arise. It is well known that acceptability judgments are constrained by performance issues, including real-world plausibility (Bader & Häussler 2010). In CS, the plausibility of a code-switch depends in part on whether or not there is a reason to code-switch. The naturalness of a code-switched utterance

will thus depend on whether there is a trigger for switching languages within the utterance (Clyne 1987). Some known triggers of CS include functional and/or discourse factors such as focus (Pfaff 1979), interlocutor (Auer 2005) and use of direct quotations (Gumperz 1982); other triggers are of a grammatical nature, such as phonology and/or prosody (McCormick 2002); and finally another trigger can be a particular lexical item (Backus 2000, 2003, McCormick 2002 and Riehl 2004). Of the many issues that can impact the perceived naturalness of a switch, lexical issues are one that is relatively easy to control. If the goal is to test whether a certain structure is acceptable or not, it is important to make sure that the choice of a particular lexical item will not interfere with the judgment. So that participants do not deem a switch unnecessary or unwarranted, as opposed to ungrammatical, it is possible to increase the naturalness of switches by using carefully selected lexical items. Consider examples (8) and (9) from the aforementioned experiment involving Taiwanese-Spanish CS.

- (8) Compró Mirta *hia-e tue-chit riab ba-tzang?*
bought M those which CL ba-tzang
'Which of those ba-tzang did Mirta buy?'
- (9) Compró Mirta *hia-e tue-chit pun tse?*
bought M those which CL book
'Which of those books did Mirta buy?'

A *ba-tzang* (8) is a rice dumpling typical of Taiwanese cuisine. It is hard to find a translation of this lexical item into Spanish, and therefore, it seems more natural to switch from Spanish into Taiwanese at this point. As a matter of fact, switches from Spanish to Taiwanese in the presence of this Taiwanese-specific lexical item were regarded as more natural by our main consultant. When asked about example (9) the consultant explicitly asked why she would switch in that context, saying that switching to Taiwanese at that point seemed unnatural to her.⁹ To put this idea in broader terms, the choice of carefully selected lexical items, e.g. lexical items that are hard to translate into the other language, can cause a switch to be perceived as more natural. Consider now another example of the importance of choosing lexical items, based on the German-Spanish sluicing study mentioned in the introduction. When designing the stimuli, the need was for verbs that assign a different case in Spanish and German. Of the verbs that satisfy this requirement, it turns out that not all of them could be used. Specifically, we avoided using some German verbs that are rare in conversational German. An example of this is *gedenken* 'to commemorate,' which assigns genitive case to its complement yet is used only in formal registers.

Looking at lexical choice in both Taiwanese-Spanish and German-Spanish studies, then, shows the importance of using carefully selected lexical items.

Further, it shows the benefits of using native code-switchers who are part of the community being studied as consultants – these speakers can help guide us to code-switches that do, indeed, sound natural. Note, however, that controlling for specific lexical items is just one way of addressing the problem. A change of topic or of interlocutor could help provide a natural trigger point as well, but in these situations it will often prove more difficult to use in intra-sentential CS. The use of phonological or prosodic triggers seems to us to also be a good way to make a switch more natural, but they can increase the complexity of the experiment, as discussed in the next subsection.

2.3.2 Modality

A final issue concerning project design is the modality in which stimuli are presented: whether stimuli are aural or written. It is known that CS can be influenced by prosody, pauses and the like (Gardner-Chloros & Edwards 2004, González-Vilbazo 2005, MacSwan 1999 and Toribio 2001a). Given that participants can invent their own context (unless a context is provided, see McCormick 2002) when making acceptability judgments (Schütze 1996), it is unclear what prosody they may be putting on written stimuli. Moreover, though CS is sometimes found in writing, it is primarily a spoken phenomenon (Grosjean 1982, Mahootian 2005 and Montes-Alcalá 2001). All of this hints at possible confounds, and while at first sight it may seem that a clear solution is to use aural (recorded) stimuli rather than written, there are pros and cons to both modalities. We can see some of the advantages and disadvantages for aural and written stimuli in (10) and (11), respectively.

(10) Advantages and disadvantages of aural stimuli

- a. Controls for phonological factors
- b. Stimuli need to be pre-tested repeatedly for phonological factors and to ensure consistency
- c. Harder to administer

(11) Advantages and disadvantages of written stimuli

- a. Easy to create and administer
- b. Does not control for phonological factors

It should be noted that bilingual linguistic competence doesn't automatically mean bilingual literacy. Code-switching is an eminently oral phenomenon and rarely found in written form. This may easily influence the judgment of code-switchers.

It is important, then, to know if the mode of stimuli presentation does, in fact, affect participants' judgments. We test this in an English-Spanish CS study on the

that-trace effect (Downey & Hoot in prep.). In this study, identical stimuli are presented in different modalities. As shown in Figure 2, the mode of stimuli presentation does, in fact, affect participants' acceptability judgments, at least for certain structures. The aural mode of presentation results in CS stimuli being rated as more acceptable across the board, with the exception of Structure 5, which was judged nearly identically across studies. In some cases, such as with Structure 7, the acceptability changes dramatically: In this example, judgments change from very questionable/ungrammatical in the written version to perfectly acceptable in the aural version, demonstrating the critical role mode of presentation can play.

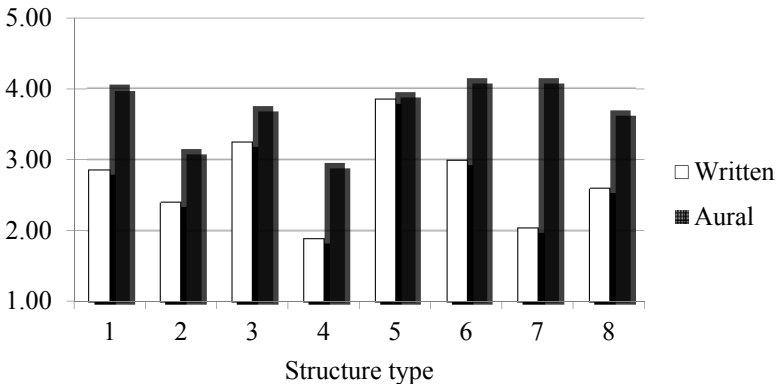


Figure 2. Mean scores of written and aural stimuli

It seems, then, that modality does make a difference in participant judgments, at least for certain stimuli types, so a conservative design would minimally pilot both modalities to see if differences do arise. However, in spite of the different scores seen in Figure 2, the choice of modality hinges additionally on weighing the importance of controlling phonological factors against the simplicity of preparing and administering a written experiment.

3. Experimental procedure

After discussing how the design of a study can isolate the relevant features, now we turn to some possible confounding variables specific to CS that can arise from within the experimental procedure used to collect data. Of course, there are many variables that need to be controlled for in experimental procedures, but here we want to address three potential confounds specific to CS: (i) the stigma attached to CS, (ii) the naturalness of the CS environment and (iii) the language mode continuum, as well as three possible solutions to these potential confounds.

It is well known that CS is often a stigmatized form of communication, particularly in certain communities (Bullock & Toribio 2009, Montes-Alcalá 2000, Poplack 1980 and Toribio 2001a). This stigma may influence and/or depress acceptability judgments, as well as cause ‘negative over-reporting’ (Henry 1992), in an effect similar to that found in non-standard language varieties in general (Henry 1995). As a result, it is difficult to know whether a particular judgment on a code-switched sentence reflects the speaker’s grammatical competence or whether the judgment might be influenced by the stigma, and this needs to be taken into consideration when eliciting judgments on CS.

Regarding the second confound, CS can be influenced by the environment in which speakers find themselves and their level of comfort within that environment (Grosjean 1998). If participants are not totally comfortable producing and/or listening to mixed language because of the artificial experimental context, their judgments could be affected. Care must be taken so that participants find CS to be natural in the experimental setting.

Finally, a third complicating factor, discussed by Grosjean (1985 et seq.), is the proposal of a continuum in the state of activation of a bilingual’s languages. At any given point in time, a bilingual’s two languages are variously activated based on contextual and interlocutive factors (Baetens Beardsmore 1986, Grosjean 1985 et seq. and Treffers-Daller 1998). This variation in language activation has the potential to affect participants’ judgments of the same stimulus – a particular code-switched sentence could be considered more or less acceptable depending on the relative activation of each language. Given this, we need to control for language activation as much as possible in the experimental setting.

To address our three confounds, we have used three potential strategies to ensure that participants’ judgments represent as much as possible their underlying competence. These solutions are outlined in Table 4.

Potential Solution	Procedural Confound		
	Stigma	Environment	Mode continuum
Instructions in CS	✗	✓	✓
Training	✓	✓	✓
Priming	✗	✗	✓

Table 4. Possible solutions to experimental procedure confounds

Our first solution, providing code-switched instructions for our experiments, is shown in (12), taken from a Spanish-*Taiwanese* CS study (Bartlett & González-Vilbazo forthcoming a, Bartlett & González-Vilbazo forthcoming b).

(12) Guan-be averiguar, si en espan-gí e-sai cambiar de español a taiwanés o de tai-gí a se-pang-ga-gí en los siguientes casos. Por favor indicá su opinión en los cuadritos, usá una escala de 1 (beh-sai kong/suena mal) a 5 (eh-sai kong/suena bien).

‘We would like to know if you can, in Spanish-Taiwanese CS, switch from Spanish to Taiwanese or from Taiwanese to Spanish in the following cases. Please indicate your opinion in the box, using a scale from 1 (sounds bad) to 5 (sounds good).’

The goal of using code-switched instructions is twofold: (i) to provide an environment where CS is clearly acceptable and (ii) to activate both of the bilinguals’ languages. By using CS in the experimental instructions themselves, we mitigate our second confound, creating an environment where CS is expected and appears natural. We also augment this, following suggestions by Grosjean (1998), by having a consultant and/or researchers present who are themselves code-switchers and, if possible, members of the community. Code-switched instructions further help to mitigate the third confound, that of potentially different levels of activation of the two languages along of the bilingual mode continuum. Reading or hearing instructions in CS necessarily activates both languages.

The second potential strategy involves training participants both to perform GJTs and to understand the task requirements in general (Culbertson & Gross 2009). This training is designed to explain the idea of a linguistic judgment using everyday language (Schütze 2005) so that participants are better able to provide judgments based on their own linguistic competence. It involves both rating a stimulus using the rating scale for the experiment and an explanation as to why the particular rating was chosen, as is shown in example (13).

(13) En cuanto al significado, the first one is a little odd. You don’t usually buy a book for a party. Pero como es posible decir esa oración, it would be rated a 5.

‘In terms of meaning, the first one is a little odd. You don’t have to usually buy a book for a party. But since it is possible to say this sentence, it would be rated a 5.’

We train first on obviously grammatical and ungrammatical monolingual utterances and then do the same with code-switched ones. Obviously, the training cannot involve any stimuli related to the experiment at hand. Note that training, just like task instructions, is done in CS to maintain the idea of a situation where CS is both accepted and expected.

Judgment task training reduces the impact of all three confounds. By showing participants that code-switched utterances can be rated as highly as monolingual

ones, we explicitly argue against the stigma that CS is somehow less acceptable than a sentence in only one language. Further, by providing training in CS and by rating code-switched examples, we continue to provide an environment where CS is natural. Finally, all training sessions include rating code-switched stimuli so that we continue to activate both of the bilinguals' languages.

The third strategy, language priming, takes advantage of the language mode continuum. Through language priming, it is possible to present bilingual participants with a task (either written or aural) that is either code-switched, thus activating both of the participants' languages, or monolingual, thus attempting to activate only one of the languages. For most of our experiments, we want participants in a bilingual mode, and therefore, priming with CS is preferred. However, additional aspects of a bilinguals' competence may be gauged by activating one language more than the other – specifically, using monolingual priming to activate a single language may make participants more likely to borrow instead of code-switch (Treffers-Daller 1998), a distinction which has been extensively discussed in the CS literature (MacSwan 1999, Myers-Scotton 2001, Romaine 1995 and Toribio 2001a). We are currently testing this solution in an English-Spanish CS study where participants judge sentences on two different days – one after being primed by CS, the other after being primed by monolingual discourse, with testing order being randomly distributed among participants (Downey in prep.). Higher ratings of ungrammatical code-switched structures after priming for CS would show positive effects of this potential solution.

The presence of a bilingual researcher or even the interest in bilingualism might prompt a bilingual mode in the subjects (Grosjean 2008). This means that any experiment that aims at putting the participants in a monolingual mode has to be designed in such a way that the participants are not aware of the researchers interest in bilingualism and obviously no bilingual researcher should be present.¹⁰

Although none of the strategies presented here is fail-safe, our experience with these strategies indicates that their use results in less variability in judgments and in a more generous use of the entire Likert scale during the GJTs. We suggest that all three solutions be employed whenever feasible so that procedural confounds can be limited as much as possible.

4. Participant selection

In addition to the confounds present within the experimental procedure, it is also important to consider the characteristics of the population being investigated, and, for code-switchers in the United States, to consider in particular the special context of bilingualism in the U.S. One issue is that bilinguals can be very different from one another. In order to ensure that this heterogeneity does not interfere with the experiment, we should consider the four characteristics of bilinguals proposed by Grosjean (1998) in (14).

- (14)a. They acquire and use their languages for different purposes, in different domains of life, with different people (Complementarity Principle, Grosjean 1997b).
- b. They are rarely equally fluent in all language skills in all their languages, as a direct consequence of the above.
 - c. They may still be in the process of acquiring a language, while others have attained a certain level of stability.
 - d. Their language repertoire may change over time with their environment and language-skills needs.

One of the ways in which we can account for some of the varying characteristics, both those listed above and others, is to test the competence of the bilingual speakers not only with respect to their CS but also with respect to both (monolingual) languages used in the CS. There are a significant number of Spanish varieties spoken in the U.S. – in part because of dialectal differences, in part because of incomplete acquisition, attrition, or acquisition of a contact variety, i.e. heritage languages (see e.g. Montrul 2010, Potowski & Carreira 2010) – as well as significant differences in the varieties of English spoken in the U.S. It is important, therefore, to consider these individual differences when studying CS. We have to establish the linguistic varieties used by a participant by investigating monolingual stimuli in both languages before analyzing CS data. Knowing the characteristics of the monolingual varieties allows us to draw from previous studies on the grammar of these varieties, and this information then can and should be used to control for the grammatical features involved in the particular CS of the experiment. Once we have the monolingual varieties, we want to group the speakers' data according to the linguistic varieties used by the participants. This does not have to take into account all possible differences but rather only the grammatical differences that might have an impact on the phenomenon under study (e.g. wh-movement, sluicing, pronouns). Failing to group participants in this way would be akin to studying ellipsis in peninsular Spanish and not controlling for the language in which the participants are competent, i.e. allowing monolingual native speakers of Argentina, Ecuador or Mexico, or even English or Mandarin, to be part of the experiment.

To illustrate, let us again revisit sluicing in German-Spanish CS. Recall that we were interested in verbs that assign a different case in German and Spanish to try to determine whether morphosyntax plays a role in licensing sluicing. One such verb pair is the Spanish *contradecir* 'to contradict' and the German *widersprechen* 'to contradict.' *Widersprechen* uncontroversially assigns dative in German. However, *contradecir* has dialectal variation in Spanish – for some speakers it assigns accusative (*lo contradije* 'I contradicted him') while for other speakers it assigns dative (*le contradije* 'I contradicted him'). In the study conducted, all bilinguals shared similar socio-economic backgrounds and were code-switchers of comparable linguistic backgrounds (they all were native speakers, went to the same school,

reported code-switching on a regular basis, etc.), yet there was nonetheless variation between speakers with regard to the case assignment of *contradecir*. Investigating their Spanish and German separately thus proved crucial to determining the case they assign from their dialect and, therefore, whether the results we predicted were warranted. By taking into account this specific linguistic variation, case assignment, we eliminated a potential confound.

However, dialectal variation need not always narrow the scope of participant selection. Again, as long as the relevant features overlap, participants of various dialectal backgrounds can be included. For instance, a study on pronouns in Spanish-English CS (González-Vilbazo & Koronkiewicz submitted) included a variety of speakers of Mexican Spanish as well as Salvadorian Spanish as participants. In this case, regardless of the dialect, the pronoun system of all participants was the same, unlike the German-Spanish CS example.

The question becomes how one identifies which participants to include. One of the easiest ways to distinguish different groups of bilinguals is through the pre-screening of relevant demographic background. As an anonymous reviewer suggests, it also could be helpful to ask participants whether they belong to a CS practicing community or not (Olson & Ortega-Llebaria 2010). Nevertheless, this is not always enough. Again, looking back at the German-Spanish CS example, case assignment is typically not present on a background questionnaire. Therefore, a more thorough way to distinguish these groups of bilinguals is via testing not only of CS stimuli but also of monolingual items. This way one can assess the features of each language, apart from CS, to ensure as much overlap within the group of participants as possible.

5. Conclusions and outlook

The purpose of this article is to highlight methodological issues in CS research and to make a modest contribution to the ongoing discussion on the matter. The usefulness of a methodology depends foremost on the object of study. In our case, that is the study of the grammatical aspects of CS, or to be more precise, the study of a grammatical phenomenon using CS data. As we laid out in the introduction, there are various methods of data gathering. Given the special circumstances of CS we opt here for a quantitative, experimental approach. It is with respect to such an approach that we address some inherent methodological problems and the strategies we have come to use to avoid or alleviate these problems. We organized these problems in three areas that we discussed independently although they are nevertheless interconnected in many ways: project design, experimental procedure and participant selection. Some of the strategies we have devised in our work we have tested independently or are currently testing independently, i.e. we have tested experimentally what impact they have on CS experiments. Some of the strategies presented here are still waiting to be tested.

Although this is not a new idea (Grosjean 1998), we believe it would be very helpful if the methodological underpinnings of a research project were reported in the literature. Ultimately, the goal should be to agree on methodological standards, both to better assess the value of reported research results and to help find better methodological tools for new research projects.

Acknowledgements

This project has been conducted under the auspices of the UIC Bilingualism Research Lab. We would like to acknowledge and thank the following people: Luis López, Kara Morgan-Short and the UIC Cognition of Second Language Acquisition Lab, Daniel Vergara, Lukasz Adamczyk, Christian Alvarado, Jesse Banwart, Blanca Bustos, Enas El-Khatib, Liz Remitz, Marlen Romero, Ivette Serrano and Jack Waas. This material is based upon work supported by the National Science Foundation under Grant No. 1146457 and by the UIC Provost's Award for Graduate Research. We would like to thank an anonymous reviewer and Timothy Gupton and the audiences of Linguistic Evidence 2012 at the University of Tübingen and the Hispanic Linguistics Symposium (HLS) 2011 in Athens/Georgia for valuable feedback on previous versions of this article. All errors remain ours.

Notes

- 1 Non-pronounced material is shown in angle brackets.
- 2 As is usual in CS research, a code-switch will be marked by having words of one of the languages in *italics*.
- 3 We do not make the claim here that all psychological studies are experimental; we are focusing here only on those of a quantitative nature.
- 4 Belazi, Rubin & Toribio 1994, DiSciullo, Muysken & Singh 1986, among others.
- 5 Examples of bilingualism-based corpora include Siarad, Patagonia, and Miami, from the Canolfan ESRC Centre.
- 6 Bartlett & González-Vilbazo forthcoming a, among others.
- 7 *Mate* is a popular tea-like beverage in Argentina.
- 8 There are some cases in which the wh-constituent can also move in Taiwanese (Craig Sailor, p.c.), but for ease of exposition we will ignore those cases here.
- 9 To paraphrase our consultant, if you are talking about books, why switch?
- 10 We would like to thank an anonymous reviewer for pointing this out to us.

References

- Auer, Peter. 2005. A postscript: Code-switching and social identity. *Journal of Pragmatics* 37, 403-410.
- Backus, Ad. 2000. The role of semantic specificity in insertional code-switching: Evidence from Dutch-Turkish. In Rodolfo Jacobson (ed.), *Code-switching worldwide*, Vol. 2, 125-154. Berlin: Mouton de Gruyter.
- Bader, Markus & Jana Häussler. 2010. Toward a model of grammaticality judgments. *Journal of Linguistics* 46, 273-330.
- Baetens Beardsmore, Hugo. 1986. *Bilingualism: Basic principles* (2nd ed.). Clevedon: Multilingual Matters.
- Bartlett, Laura & Kay González-Vilbazo. Forthcoming a. Classifiers in Spanish-Taiwanese code-switching. In *Proceedings of the Chicago Linguistic Society (CLS) 46*.
- Bartlett, Laura & Kay González-Vilbazo. Forthcoming b. The structure of the Taiwanese DP: Evidence from Taiwanese/Spanish code-switching. *Journal of East Asian Linguistics* 22.
- Belazi, Hedi M., Edward J. Rubin & Almeida Jacqueline Toribio. 1994. Code-switching and X-Bar Theory: The functional head constraint. *Linguistic Inquiry* 25, 221-237.
- Bullock, Barbara E. & Almeida Jacqueline Toribio. 2009. Themes in the study of code-switching. In Barbara Bullock & Almeida Jacqueline Toribio (eds.), *The Cambridge handbook of linguistic code-switching*, 1-17. Cambridge: CUP.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, Noam. 1986. *Knowledge of language*. New York: Praeger.
- Chung, Sandra. 2006. Sluicing and the lexicon: The point of no return. In Rebecca T. Cover & Yuni Kim (eds.), *Proceedings of the annual meeting of the Berkeley Linguistics Society* 31, 73-91. Berkeley: Berkeley Linguistics Society.
- Clyne, Michael. 1987. Constraints on code-switching: How universal are they? *Linguistics* 25, 739-764.
- van Craenenbroeck, Jeroen. 2010. *The syntax of ellipsis: Evidence from Dutch dialects*. New York: Oxford University Press.
- Culbertson, Jennifer & Gross, Steven. 2009. Are linguists better subjects? *British Journal for the Philosophy of Science* 60, 721-736.
- Di Sciullo, Anne-Marie, Pieter Muysken & Rajendra Singh. 1986. Government and code-mixing. *Journal of Linguistics* 22, 1-24.
- Downey, Sarah. In prep. Gender agreement in bilingual borrowing.
- Downey, Sarah & Brad Hoot. In prep. Methodology in code-switching.
- Gardner-Chloros, Penelope & Malcolm Edwards. 2004. Assumptions behind grammatical approaches to code-switching: When the blueprint is a red herring. *Transactions of the Philological Society* 102, 103-129.

- Genesee, Fred. 1989. Early bilingual development: One language or two? *Journal of Child Language* 16, 161-179.
- González-Vilbazo, Kay. 2005. *Die Syntax des Code-Switching*. Doctoral dissertation, University of Cologne.
- González-Vilbazo, Kay, Laura Bartlett, Shane Ebert & Daniel Vergara. In prep. Wh-phenomena in Taiwanese-Spanish code-switching.
- González-Vilbazo, Kay & Bryan Koronkiewicz. Submitted. Pronouns in Spanish-English code-switching.
- González-Vilbazo, Kay & Luis López. 2012. Little *v* and parametric variation. *Natural Language & Linguistic Theory* 30, 33-77.
- González-Vilbazo, Kay & Sergio E. Ramos. Forthcoming. A morphosyntactic condition on sluicing: Evidence from Spanish/German code-switching. *Linguistic Inquiry*.
- Grosjean, Francois. 1982. *Life with two languages: An introduction to bilingualism*. Cambridge: Harvard University Press.
- Grosjean, Francois. 1985. The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development* 6, 467-477.
- Grosjean, Francois. 1997a. Processing mixed language: Issues, findings, and models. In Annette M. B. de Groot & Judith F. Kroll (eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, 225-254. Mahwah: Lawrence Erlbaum Associates.
- Grosjean, Francois. 1997b. The bilingual individual. *Interpreting* 2, 163-187.
- Grosjean, Francois. 1998. Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition* 1, 131-149.
- Grosjean, Francois. 1999. Individual bilingualism. In Bernard Polsky (ed.), *Concise encyclopedia of educational language*. Oxford: Elsevier.
- Grosjean, Francois. 2001. The bilingual's language modes. In Janet L. Nicol (ed.), *One mind, two languages: Bilingual language processing*, 1-22. Oxford: Blackwell.
- Grosjean, Francois. 2008. *Studying bilinguals*. Oxford: Oxford University Press.
- Gullberg, Marianne, Peter Indefrey & Pieter Muysken. 2009. Research techniques for the study of code-switching. In Barbara Bullock & Almeida Jacqueline Toribio (eds.), *The Cambridge handbook of linguistic code-switching*, 1-17. Cambridge: CUP.
- Gumperz, John. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.
- Henry, Alison. 1992. Infinitives in a for-to dialect. *Natural Language and Linguistic Theory* 10, 279-301.
- Henry, Alison. 1995. *Belfast English and standard English: Dialect variation and parameter setting*. New York: Oxford University Press.
- MacSwan, Jeff. 1999. *A minimalist approach to intrasentential code switching*. New York: Garland Pub.

- Mahootian, Shahrzad. 2005. Linguistic change and social meaning: Code-switching in the media. *International Journal of Bilingualism* 9, 361-375.
- McCormick, Kay. 2002. Code-switching, mixing and convergence in Cape Town. In Rajend Mesthrie (ed.), *Language in South Africa*, 216-234. Cambridge: Cambridge University Press.
- Meisel, Jürgen M. 1990. Two first languages: Early grammatical development in bilingual children. Dordrecht: Foris.
- Merchant, Jason. 2001. The syntax of silence: Sluicing, islands, and the theory of ellipsis. Oxford: Oxford University Press.
- Merchant, Jason. 2008. An asymmetry in voice mismatches in VP-ellipsis and pseudogapping. *Linguistic Inquiry* 39, 169-178.
- Montes-Alcalá, Cecilia. 2000. Written code-switching: Powerful bilingual images. In Rodolfo Jacobson (ed.), *Code-switching worldwide*, Vol. 2, 59-74. Berlin: Mouton de Gruyter.
- Montrul, Silvina. 2010. Current issues in heritage language acquisition. *Annual Review of Applied Linguistics* 30, 3-23.
- Myers-Scotton, Carol. 2001. The matrix language frame model: Developments and responses. In Rodolfo Jacobson (ed.), *Code-switching worldwide*, Vol. 2, 23-58. Berlin: Mouton de Gruyter.
- Olson, Daniel and Marta Ortega-Llebaria. 2010. The perceptual relevance of code switching and intonation in creating narrow focus. In Marta Ortega-Llebaria (ed.) *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, 57-68. Somerville, MA: Cascadilla Proceedings Project
- Pfaff, Carol W. 1979. Constraints on language mixing: Intra-sentential code-switching and borrowing in Spanish/English. *Language* 55, 291-318.
- Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish: Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics* 18, 581-618.
- Potowski, Kim & Maria Carreira. 2010. Spanish in the USA. In Kim Potowski (ed.), *Language diversity in the USA*, 66-80. Cambridge: Cambridge University Press.
- Riehl, Claudia Maria. 2004. Code-switching in bilinguals: Impacts of mental processes and language. In James Cohen, Kara T. McAlister, Kellie Rolstad & Jeff MacSwan (ed.), *ISB4: Proceedings of the 4th International Symposium on Bilingualism, 1945-1959*. Sommerville, MA: Cascadilla.
- Romaine, Suzanne. 1995. *Bilingualism*. Oxford: Blackwell Publishers Ltd.
- Ross, John Robert. 1969. Guess who. In Robert I. Binnick, Alice Davison, Georgia M. Green & Jerry L. Morgan (ed.), *Proceedings of the fifth regional meeting of the Chicago Linguistic Society*, 252-286. Chicago: The University of Chicago Press.
- Schütze, Carson T. 1996. *The empirical base of linguistics*. Chicago: The University of Chicago Press.

- Schütze, Carson T. 2005. Thinking about what we are asking speakers to do. In Stephen Kepsler & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 457-484. Berlin: Mouton de Gruyter.
- Toribio, Almeida Jacqueline. 2001a. Accessing bilingual code-switching competence. *The International Journal of Bilingualism* 5, 403-436.
- Toribio, Almeida Jacqueline. 2001b. On the emergence of code-switching competence. *Bilingualism: Language and Cognition* 4, 203-231.
- Treffers-Daller, Jeanine. 1998. Variability in code-switching styles: Turkish-German code-switching patterns. In Rodolfo Jacobson (ed.), *Code-switching worldwide*, 177-197. Berlin: Mouton de Gruyter.
- Zentella, Ana Celia. 1997. *Growing up bilingual*. Malden: Blackwell.

Copyright of Studies in Hispanic & Lusophone Linguistics is the property of Studies in Hispanic & Lusophone Linguistics and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.