Structural Features and Sociolinguistic Functions in Bilinguals' Code Switched Face-to-Face and Short Message Service Conversations

Susana M. Sotillo, Ph.D.
Department of Linguistics
Montclair State University
Montclair, NJ 07043-1624, USA.

Abstract

This study investigates morph syntactic features and sociolinguistic functions of code-switching in face-to-face (F2F) and short message service (SMS) conversations among bilinguals from typologically different languages: Spanish and Tagalog. Intra sentential code-switching (ICS) is first analyzed in four bilingual corpora. Next, conversation topics in sequential exchanges are examined to infer sociolinguistic functions. Embedded language islands (ELIs) were more frequent in SMS than in direct conversations because more processing time for selecting lemmas is available to those texting than to individuals engaged in F2F interaction. However, inserts were more frequent in F2F conversations. The inferred sociolinguistic functions of CS among bilinguals include arguing, criticizing, gossiping, negotiating message comprehension, and relaying information. These findings reveal the inherent variation of ICS practices in different modes of communication and show how bilinguals in interaction creatively exploit their varied semiotic resources of which language is one

Key Words: Bilinguals, Code-switching, Conversations, Morph syntax, SMS, Sociolinguistics, Spanish, Tagalog.

Code-switching (hereafter CS) or the alternation between two codes or languages within a stretch of discourse is a global phenomenon. The focus has often been on either inter sentential CS which occurs at the boundaries of clauses or sentences, or intra sentential CS (ICS), present within a clause or sentence (Gumperz, 1982; Poplack, 1980, 1988; Woolard, 2004; Zentella, 1982, 1997). CS is also a frequent feature of digital or computer-mediated communication (CMC), a hybrid form of immediate and delayed communication that includes texting or Short Message Service (SMS). This study investigates the morph syntactic constraints and sociolinguistic functions of ICS in face-to-face (F2F) and short-message service (SMS) conversations among Spanish-English and Tagalog-English bilinguals. It employs a cognitive approach based on the Matrix Language Frame(MLF)model and the 4-M model for the structural analysis of four bilingual corpora which allows researchers to go beyond surface descriptions of CS. Wei's Bilingual Lemma Activation model (BLA) and Androutsopoulos' (2007) approach to CS in digital communication also inform this study.

1. Limitations of Previous Investigations

Poplack's studies from a Variationist perspective (1980, 1988) using traditional grammatical categories did not account for certain prepositions which Myers-Scotton (1993) and Myers-Scotton & Jake (2000, 2009) classify as content morphemes because they assign thematic roles. CS based on Conversational Analysis (CA) and Social Network Analysis also informed previous studies (Auer, 1988, 1995; Milroy, 1987), but CA has its shortcomings since it does not have a fixed theory and focuses on sequentially organized speech behavior (Gafaranga, 2009).The present study uses a combination of qualitative and quantitative approaches to investigate and compare ICS in both F2F and SMS conversations so as to identify patterns and explain findings. Code-switched topics in sequential exchanges are also analyzed in order to infer sociolinguistic functions.

1.1 CS in Contemporary Digital Communication

The evolution of CMC on a global scale and the expansion of language choice in multilingual web-based discussions, multicultural forums, and discussion lists have been extensively investigated (Danet & Herring, 2007; Paolillo, 2007).

Other researchers have shown that young people manage technological constraints in digital communication and use a 'glocal' language that indexes a group's ethnicity or lifestyle choices (e.g. Androutsopoulos, 2007; Anis, 2007; Chiluwa, 2008; Dürschei & Stark, 2011; Simango, 2011).Marzuki (2013) found that apologies sent via SMS in Malay differed from those sent via SMS in English by the same participants, with higher frequencies of vowel/consonant omissions, use of colloquialisms, and rebus abbreviations sent in Malay. Likewise, Anis (2007) showed that the language of SMS deviated from the prescriptive norms of standard French and regarded such CMC language output "as the interaction of technical, economic, communicative, and psychosocial constraints" (p. 110).

1.2 SMS or Texting Communication

Few journal articles have published on language alternation in SMS discourse. For example, Chiluwa (2008) examined patterns of CS in Nigeria. In South Africa, Deumert & Masinyana (2008) investigated CS in English and isiXhosa in stretches of text messages between and among young adults. They found that the texts in isiXhosa were using two non-overlapping sets of sociolinguistic norms. Fairon, Klein, & Paumier (2006) found that linguistic creativity and modes of expression were not very different from those of standard French and called for a more careful redefining of the differences between written and spoken language. Morel, Bucher, Doehler, & Siebenhaar (2012) uncovered neologisms and semantic shift, pseudo-borrowings, homographs/homophones and ideographic switching in SMSs from multilinguals and urged a broadening of the concept of CS.

The present study investigates the structure of ICS in both F2F and SMS communication, and the sociolinguistic functions accomplished through the CS practices of Spanish-English and Tagalog-English bilinguals. Code switched SMS corpora were downloaded automatically with software for specific Smart phones such as Androids and iPhones, and annotated transcripts of F2F conversations and interviews were obtained from large databases available online (the Miami Bangor corpus and ICE Philippines corpus).The theoretical framework and categories from Myers-Scotton & Jake's (2000, 2009) MLF model, 4-M model, and Wei's (2009) BLA Model were used in a detailed analysis of ICS in F2F and SMS conversations. I argue that the theoretical import and rigorous analysis of linguistic data that inform these approaches will help explain what goes on in the mind of proficient and fluent bilinguals of the two languages examined in this study.

2. Theoretical Framework

The 4-M model classifies morphemes into four types based on their election at the conceptual level. It further specifies that information concerning content and early system morphemes (i.e., system morphemes that occur with content morpheme heads) is salient at the conceptual level, whereas information concerning late system morphemes becomes salient only at the positional level of the Formulator. Wei's (2009) BLA seeks to describe and explain the cognitive basis of ICS by asserting that ICS cannot be accounted for by descriptive observations of surface phenomena, but that this phenomenon needs to be explained as "cognitively based operations of an abstract nature" (Wei, 2006: 149).This study argues that similar procedures are also theorized to take place in instances of ICS in SMS communication.

2.1 Matrix vs. Embedded Languages

The Matrix Language (ML) or base language, and the Embedded Language (EL) constitute key features of ICS research, with the ML playing a dominant role because it can provide the grammatical or sentential frame for the bilingual clause (Wei, 2006). This key distinction accounts for the differential participation of the ML and the EL in influencing ICS utterances or texting, as is the case in this study. A match between the ML and the EL in ICS at the lemma level is known as lemma congruence. The ML controls morph syntactic procedures and supplies most of the content and system morphemes (Wei, 2006, 2009).

2.2 Content vs. System Morphemes and Compromise Strategies

The 4-M Model categorizes system morphemes (SMs) in three ways: early system morphemes and two late system morphemes: bridges and outsiders (Myers-Scotton & Jake, 2000).Definite articles, plural and derivational affixes are classified as early system morphemes. The opposite is the case with two types of late SMs: bridges and outsiders. Bridges such as English "of" and the possessive ('s) connect elements that are part of larger constituents. An outsider morpheme "depends on information that is outside of the element with which it occurs," such as subject-verb agreement and the suffixes in pro-drop languages (Myers-Scotton & Jake, 2009: 346).

Lexical representations are language specific because languages lexicalize concepts differently; thus, when lemmas or abstract representations activated from the EL do not sufficiently match the ML counter parts, a compromise strategy is chosen, such as the production of ELIs in ICS (e.g., LET ME KNO Wsinopido un RIDE ami amiga–PA, 19, F). ELIs consists of an EL content morpheme with only other EL morphemes, including system morphemes.

Two research questions are addressed: 1. Are the types of morphemes in ICS and shifts in ML found in F2F conversations different from those in SMS or texting? In addition, 2. What sociolinguistic functions are evident in these two different modes of communication?

3. Methodology

3.1 SMS Data Collection and Participants

The Spanish-English SMS corpus consisted of 2338 words and 789 types. These were conversations between an SMS social network owner and 82 different individuals. The SMSs were contributed by a graduating senior and active texter who spoke Peruvian Spanish. Among those sending messages to him were speakers of Colombian, Ecuadorian, and Argentinian Spanish. All interactants engaged in CS when texting. Messages ranged in word length from three to 53words. SMS participants ranged in age from 18 to 56, with an average age of 23, and included39 females and 43 males. In contrast, the SMS Tagalog-English corpus consisted of 3010 words and 930 types donated to the researcher's ongoing SMS data collection project of CS in multiple languages. The owner of this SMS social network was an ESL teacher at a local northern New Jersey community college. Those sending text messages to her were primarily colleagues, friends, and family members. Messages that were part of 20 different conversations varied in length from one word plus an emoticon to 27words.A total of 13 participants, three males and 10 females who ranged in age from 21 to 65, with an average age of 37, formed part of this Tagalog-English bilingual's SMS social network.

3.2 F2F Data Collection and Participants

Thirty-nine files (transcripts) of F2F conversations out of 56 were randomly selected and downloaded from the Talk bank Bangor-Miami corpus (https://talkbank.org/). The transcript consisted of 19174 words and3457 word types. There were 1099 instances of Spanish-English and English-Spanish ICS, 4995 intersentential switches, and 13080 instances of no CS in the conversations analyzed. The F2F Tagalog-English data were based on 22 transcribed conversations out of 300 spoken files from the International Corpus of English (ICE) Philippines Corpus (http://ice-corpora.net/ice/icephi.htm), which is downloadable from the ICE website. The samples analyzed consisted of 7477 word tokens and 1399 word types.

3.2 Procedures

3.2.1Coding in F2F and SMS Corpora

It was predicted that quantitative and qualitative differences in morph syntactic features in ICS in these two modes of communication could be explained from a cognitive perspective, whereas the digitally contextualized nature of SMS would account for differences in sociolinguistic functions. Different types of morphemes, ELIs, and inserts, consisting of tags, interjections, expletives, discourse markers, etc., as classified by Biber et al. (2002: 449-454), were identified and manually coded by the researcher and her team. Following the manual coding, the team tagged all morphemes based on the 4-M model of Myers-Scotton & Jake (2000, 2009) with a slight modification by including inserts as a category in all four corpora. A parser was designed to process instances of ICS. These were extracted from 39 files from the Miami Corpus, each containing over a thousand words, numbers, and symbols. Since language markers were present in the data set such as "@s: spa" or "@s:eng," instances of ICS were extracted by obtaining lines of the data set that contained combinations of language markers. Regular expressions were used to remove the remaining noise from the data set in question. The parser was also used to process instances of ICS in 22 files of transcribed conversations among college students, friends, neighbors, and professionals obtained from the ICE Philippines English corpus.

Working separately on the Spanish-English transcribed conversations and SMS data, examples of various types of morphemes classified according to the 4-M Model in ICS episodes were first identified, manually coded, and subsequently tagged by the principal investigator and a graduate assistant. Bilingual undergraduates proficient in Tagalog-English were trained to identify, code, and tag all morpheme types, ELIs, and inserts in the ICE Philippines and Tagalog-English SMS corpora.

Coded transcripts by the researcher and her team based on the Spanish-English and Tagalog-English corpora were submitted to an expert in bilingualism to check for accuracy. Interrater reliability coefficients obtained using SPSS for coding accuracy in the Spanish-English and Tagalog-English F2F corpora were .947 and .851, respectively. For the Spanish-English and Tagalog-English SMS corpora, interrater reliability coefficients were .92 and .82, respectively.

The research questions addressed the types of morphemes, ELIs, and inserts present in ICS in different types of corpora, whether or not MLs shifted during ICS, and the possible functions or purpose of CS in F2F and SMS contexts. Pearson Chi-Square values and Cramer's V (symmetric measures) were calculated in order to determine whether the observed vs. expected frequencies for morphemes, ELIs, and inserts found in ICS in two different modes of communication were significant. Since the F2F Spanish-English and Tagalog-English corpora were larger than the SMS Spanish-English and Tagalog-English corpora, raw frequencies for all four corpora were normalized (see Biber, Conrad, & Reppen, 1998: 263).

3.2.2 Topics of Conversation and Functions of CS

There search team read and analyzed all conversations in the Spanish-English corpora following Androutsopoulos' (2007) procedures from his study of German-based online diasporic forums. A fluent and proficient Tagalog-English bilingual read and identified the topic/theme of conversations in the Tagalog-English corpora. We inferred the functions of CS in these two different modes of communication. Those who contributed SMSs and their social network participants were asked to complete a short online survey explaining their reasons for CS when sending text messages.

The following section presents the results of the morph syntactic analysis in instances of ICS that were found in these corpora. Examples of various types of morphemes, ELIs, and inserts were randomly selected from the beginning, middle, and end of transcribed F2F conversations. A similar procedure was followed for selecting examples from the SMS conversations.

4. Results

4.1 Analysis of Spanish-English Corpora

Speakers activate language-specific lemmas in the BML based on their pre-verbal communicative intention. The following are examples of content morphemes from the Miami corpus activated during ICS when the sentential frame is Spanish and then following a shift to a sentential frame in English:

Spanish-English

(1)   vas al GROCERY coño (content morpheme followed by common insert/expletive)
        [You]2.SING are going to-V-ING **nt-EXPLETIVE?
        "Are you going to the grocery store **nt?"
(2)   *esoes* HORSE COUNTRY
        That-DEM is-VPRESNM
        "That is horse country."

English-Spanish

(3)   IT'S COLOMBIAN AND THEIR *empanadas*
        DET pies- N
        "It's Colombian and their pies"
(4)   *Pero* SHE HAS LIKE *una mala fama*
        But-CONJ   a-DETbad-AD Jreputation-N
        "But she has a bad reputation."
(5)   *que* THE PRIMARY PURPOSES OF THE EDUCATORS IS THE STUDENTS
        that COMP
        "that the primary purpose of educators is the students."

Example (1) shows a content morpheme in English, "grocery," embedded within the sentential frame of Spanish. Example (2) illustrates how English language content morphemes are embedded within Spanish sentential frame and do not violate the syntax of either English or Spanish.

The word order of the EL, English, is preserved as in "horse country," with "horse" modifying "country."In (3) like comosifueraunamadre (like/as a mother), "like" is an adverb modifying an entire utterance. It is classified as a content morpheme according to the 4-M Model, which differs considerably from other traditional models that classify morphemes based on major lexical categories and simply count open- and closed-class items instead of explaining how the bilingual mind accesses lemmas directly at the level of the Conceptualizer before sending directions to the Formulator for speech production.

When switching ML frameworks, as in (3) through (5), content morphemes are switched in to an English-language sentential frame (3) as "empanadas" and (4) "una mala fama."Certainsystem morphemes from the EL, Spanish, may play a special grammatical role as in (4) "pero," and (5) "que," since they actually introduce the English ML frame work. All these content and system morphemes have been examined and coded within the context of F2F conversations.

The following are examples of ICS from SMS or text messages which are instantiated as either ELIs in English or bridges (late system morphemes) introducing clauses or sentences:
Spanish-English

(6)   Le di el WRONG NUMBER   (JA, 20, M)
        To him-IND OBJ PROI-1.SING gave-VPAST the wrong number.
        "I gave him the wrong number."
(7)   Si, era mi FINAL PROJECT   (MC, 21, M)
        Yes-AFFIRM was-VPAST my-POSS
        "Yes, it was my final project."
(8)   Tengoque irtemprano…BUT I DON'T KNOW IF IWANNA GOBECAUSE I GOT   NOTHING TO
        DO LATER   (MK, 20, M)
        [I-1.SING]] have to go-OBLIG INFearly-ADV
        "I have to go early but I don't know if I wanna go because I got nothingtodo later."
(9)   Mira yatetoy FREKEANDO AHAHA(GG, 24, M)
        Look-ATT GETTERI am-1.SING PRES freaking out PROG-ING you-ACC
        "Look, I am freaking youout, ahaha"
(10)   Tengoque arreglar MY COUSIN'S COMPUTER   (JA, 25, M)
        [I-1.SING] have to-OBLIG INF fix-VPRES
        "I have to fixmy cousin's computer."

By examining the context of the above ICS exchanges, we were able to determine that Spanish provides the sentential framework. In examples (6) through (10), ICS instances from SMSs, the embedded content morphemes follow the structure of English, with adjectives modifying nouns as "in wrong number" and "final project." Bridge system morphemes such as BUT in example (8) occurs at clause boundaries as both the sender and recipient of an SMS switch to English. BUT, as a system morpheme, is part of the EL sentence that follows, which is congruent with the framework of the ML and represents and economical way of conveying message meaning. In (9), "freak" is an EL content morpheme from English switched into the Spanish ML grammatical frame. he verb "estoy" (to be) has been shortened as "toy," and the suffix –"ando" work together well to formulate the Spanish progressive aspect. Thus, "estoy" has been shortened for ease and speed of texting. In (10), my cousin's computer retains the NP possessive construction in English, whereas Spanish requires the obligatory preposition "de" (la computadora de mi primo).

In the examples below, "pero" (but) in (11) is used in sentence initial position connecting a response to a previous excuse given by a texter for not going to a house party. In (12), the conjunction "que," a system morpheme, is found in clause initial position as a texter switches from Spanish to English when describing a party guest."Porque" (because) in(13) is used in response to a previous message askingan interlocutor for an explanation for a statement made (¿porqué?or why?):"porque (because) it's not gonna be packed."In (14), "que" is a content morpheme, standing for "que tecayo mal" (that which made you sick), a relative pronoun in an embedded attributive Spanish language clause:

Switches from Spanish to English

(11)   pero THAT'S NOT A BIG ISSUE, A …
        But-CONJ

"but that's not a big issue, a…" (PL, 24, M)

(12) toysegura que TINA WOULD BE A PRETTY GIRL[I]am-1.SING PRES sure-ADJ that-CP
"I am sure that Tina would be a pretty girl"   (JA, 25, M)

(13) Porque IT'S NOT GONNA BE PACKED
Because-SUB CONJ   (PL, 21, M)
"because it's not going to be packed" (adverbial-like subordinator)

English-Spanish

(14) WHAT DID YOU EAT YESTERDAY que tecayo mal? (PA, 19, F)
That-COMP   you-EXPERIENCER   made-VPAST   ill-ADJATTRIBUTIVE   CLAUSE   (ELI Spanish)
"What didyou eat yesterday that made you sick?"

The ELI in (14) is congruent with the sentential framework provided by the ML (English), and does not violate the linguistic economy of a code-switched text message. There is a switch in (12) from a Spanish clause, "estoysegura que" (I am sure that…), to an ELI in English (N + modal + BEV + ADJ Phrase) that has a complementizer from Spanish.

4.2 Analysis of Tagalog-English Corpora

The principal languages in the Philippines are Tagalog, Filipino, and English. Filipino is a Tagalog-based language which includes some words from regional and local dialects like Cebuano, Ilokano, Kapampangan, Hiligaynon, Bicolano, and Waray. Comprehensive linguistic and sociolinguistic studies of Tagalog-English CS were initially carried out by (Bautista, 1990, 2004). An important finding is that CS is a fact of life in all social classes and domains in the Philippines (Rafael, 1995; Smedley, 2013).Constructions in Tagalog usually follow aV-S-O/S-O-V word order, as in Kumain (eat) ako (I) ng(a/the) saging (banana), which differs from the canonical S-V-O word order of English, and passive constructions show internal changes (infixes)in the verb.

The following are examples from English-Tagalog, where the ML was identified as English:

English-Tagalog:

(15) NO SHOCK parang I <.> a </.> kaya ngaI ASKED HER YESTERDAY RIGHT
Like-ADV   so-CONJ so-CONJ
"No shock, like, so I asked her yesterday, right?" (System Morpheme - ELI)

(16) YEAH HE'S REALLY YOUNG   *pa, di ba*
still-ADV, Right?-TAG/CONJ QUES AFFIX
"Yeah he's really young still, right/no?"(Content morpheme and insert)

(17) Hay nakuTrixie ALSO UHMOh-my goodness-EXPLETIVE
"Ohmy goodness, Trixie also, uhm" (Hay naku is an expression that simply expresses   exasperation on the part of the speaker.)

(18) WE WERE HUGGING EACH OTHER 'CAUSE IT ITWAS LIKE LIKE anglakasgrabe
It-DETlarge-ADJ excessive-ADJ
"We were hugging each other, 'because it was really a strong [earthquake]." (ELI)

(19) O sige BUY ulam
Ohgo [ahead] EXCL INTJfood-NOUN (Content)
"Oh, go ahead, buy food" – (Inserts-content morpheme)

Although in (19) Tagalog lexemes predominate, the word order is definitely English.
In the examples below, Tagalog is the ML, but many of the inserts are vestiges from Spanish:
Tagalog-English:

(20) SUPER ganoonlang EXAGGERATE
That-PRO just-ADV PART
"That's just super exaggerated" – (EL Content morphemes)

(21) Oynaririnigniyodaw FUTURE WIFE niDENS
[i]PRO hear-PASTyou-2.PRO   said-V Den's-SING POSS MARKER
"I heard yousay Den's future wife..." – (EL Content morphemes)

(22) Basta STRICTLY FRIENDS lang

16

    Enough-INTJ   only-ADV

    "That's it, strictly friends, only" – (Insert and Content morpheme)

(23)   ESPECIALLY yungmgamgaanoyungmga TRYING TIMES

    That-CONJthose-PL DET what-DET what-DET those PL DET [are-[ay]V]

    "Especially that these are trying times." – (Content morphemes)

The following tagged lines are from the SMS corpus of a Tagalog-English bilingual. English as the MLis illustrated in the following examples:

(24)   OF COURSE syempre<CM>TRADITIONAL</CM>ako, HEHEHE<CM>Ano</CM>

    of course-INSERT/AFF   I-PRON + am-V(BE)

    What-DET PRON +V(BE) DETCOSTUME<CM>mo</CM>?

                           You-POSS PRO

    "Of course. I'mvery traditional. What's your costume?"

(25)   <CM>Ako</CM><CM>HERE</CM>here<CM>na</CM><CM>sa</CM>school.

    (AU, 36, F)

    I-PRO + VBE   to-PREP   to-PREP

    "I'malso here now in school."

Though Tagalog morphemes are used in (25),the entire transcribed utterance follows the syntactic structure of English. This becomes clear when all lexemes are translated from Tagalog to English. Thus, English is identified as the ML.

In all instances of ICS in the Philippines corpus, English was identified as the ML 88% of the time, whereas the SMS Tagalog-English corpus was more evenly divided with respect to the dominant ML. The results of the structural analysis of linguistic data in F2F and SMS conversations are displayed in Tables 1 through 4 as raw frequencies for different types of morphemes (content, system, ELIs, and inserts).

Table 1: Morpheme Types in F2F ICS – Talk bank-Bangor Miami Corpus

Matrix Language as % of ICS instances	Content Morphemes	System Morphemes	Embedded Language Islands (ELIs)	Inserts	Totals
Spanish (49%)	162	80	86	214	542
English 51%)	202	75	82	198	557
Totals	364	155	168	412	1099

As shown in Table 1, in the sample analyzed inserts (412) occurred frequently in the F2F Miami corpus, followed by content morphemes (364), ELIs (168), and system morphemes (155).The ML alternated as individuals switched languages back and forth, with English being the ML 51%of the time and Spanish 49% of the time. Frequencies for different types of morphemes present in the corpus of Spanish-English SMS conversations are displayed in Table 2.

Table 2: Morpheme Types in Spanish-English ICS - SMS Corpus

Matrix Language as % of ICS instances	Content Morphemes	System Morphemes	Embedded Language Islands (ELIs)	Inserts	Totals
Spanish (49%)	23	13	33	10	79
English 51%)	28	17	36	9	90
Totals	51	30	69	19	169

In the Spanish-English SMS corpus, ELIs (69) occurred to a greater extent than other types of morphemes, followed by content morphemes (51), system morphemes (30), and inserts (19).There were 499 instances of CS in the Spanish-English SMS corpus; however, whereas inter sentential switches accounted for 66% of all CS, ICS accounted for only 34%. In other words, most of the CS in the Spanish-English SMS corpus whose owner spoke a Peruvian variety of Spanish occurred at sentence boundaries and not primarily within phrases or clauses. English was the ML 51% of the time and Spanish 49% of the time. In F2F conversations from the ICE Philippines Corpus selected for analysis, 365 instances of CS were examined. ICS accounted for over 80% of all CS and a random sample of ICS instances was chosen for analysis. ICS percentages were higher in the Philippines corpus than percentages of ICS present in the Miami corpus of F2F conversations. Raw frequency counts for different types of morphemes, including ELIs and inserts, are displayed in Table 3. Tagalog-English Concurred primarily within phrases and clauses.

Table 3: Morpheme Types in F2F ICS – Philippines Corpus

Matrix Language as % of ICS	Content Morphemes	System Morphemes	Embedded Language Islands (ELIs)	Inserts	Total
Tagalog (12%)	13	4	3	2	22
English (88%)	98	18	5	39	160
Total	111	22	8	41	182

When Tagalog was identified as the ML in ICS, content morphemes (13), which included early system morphemes, were more frequent han all other types of morphemes examined in the Philippines corpus conversations. These include dearly system morphemes identified in the corpus, "which typically occur with content morpheme heads that select them" (Myers-Scotton & Jake, 2009: 342).For example, most prepositions, verb particles, definite articles occur with NPs. They can be content morphemes as in "he walked across the street," where "across" assigns a thematic role. Bridges and outsiders as part of late system morphemes were also identified in F2F Tagalog-English ICS. Frequencies for other types of morphemes with Tagalog as the ML were as follows: System morphemes (4), ELIs (3), and ITIEs (2).The system morphemes identified were primarily late system morphemes. Tagalog was the ML in 12% of all ICS, whereas English provided the sentential frame 88% of the time.ICS episodes in conversations where English was the ML had a higher frequency of content morphemes (98), followed by inserts (39) and system morphemes (18).There were very few ELIs (5).

As the following examples show, the ML switches back and forth between Tagalog and English in the Philippines corpus:

(26)   Tagalog ML:CM>Ano</CM><SM>ba</SM><CM>naman</CM>
    <CM>yung</CM> every month she'll save up

(27)   English ML:Uh <SM>kasi</SM><CM>siya</CM><CM>rin</CM>as of now he's because he also because he is also planning to put up a business<CM>pero</CM>

Table 4: Morpheme Types in ICS – Tagalog-English SMS Corpus

Matrix Language as % of ICS Instances	Content Morphemes	System Morphemes	Embedded Language Islands (ELIs)	Inserts	Total
Tagalog as ML (55%)	66	2	26	0	94
English as ML (45%)	45	21	2	10	78
	111	23	28	10	172

As shown in Table 4, during ICS episodes in these SMS conversations the MLs shift back and forth, with English being the ML 45% of the time and Tagalog 55% of the time. When Tagalogis the ML, content morphemes (66) are switched more frequently than any other types of morphemes. ELIs (26) is more frequent when Tagalog is the ML in SMS conversations than when English is the ML (2).No inserts switched into English were identified when Tagalog was the ML as shown above. When English was the ML, content morphemes (45) and system morphemes (21) were more frequent than inserts (10) or ELIs (2).

5. Significance Tests

The results of a Chi-square test show that there were indeed significant differences between the observed and expected frequencies in ICS for different types of morphemes by mode of interaction; F2F vs. SMS. Table 5 displays the Chi-Square results.

Table 5: Observed and Expected Frequencies of Association between Modes of Interaction and Morpheme Types in Spanish-English ICS

Morpheme Types	Mode of Interaction	
	F2F	SMS
Content	364 (359.7)	51 (55.3)
System	155 (160.3)	30 (24.7)
ELIs	168 (205.4)	69 (31.6)
Inserts	412 (373.6)	19 (57.4)

$\chi2$ 82.534
df   3   Sig.(2-sided)
p   <.0001
Cramer's V *.255*

With respect to the two interactional modes, F2F and SMS, it seems that the frequencies of system and content morphemes do not differ much between them. Most of the contribution to the fairly large Chi-Square comes from the differences in ELIs and inserts; ELIs or compromise strategies are more likely to occur in the SMS mode and less likely to occur in the F2F mode, where as inserts are more likely to occur in the F2F mode but less likely to occur in the SMS mode.

Table 6: Observed and Expected Frequencies of Association between Modes of Interaction and Morpheme Types in Tagalog-English ICS

Morpheme Types	Mode of Interaction	
	F2F	SMS
Content	111(114.1)	111(107.9)
System	22(23.1)	23 (21.9)
ELIs	8 (18.5)	28 (17.5)
ITIEs	41 (26.2)	10(24.8)

$\chi2$ 29.718
df   *3,*   Sig.(2-sided)
p   <.0001
Cramer's V .290

The results in Table 6 show that the overall test is significant at the p<.0001 level. These results are similar to those obtained for the observed and expected frequencies of association between two different modes of communication in the Spanish-English corpora. System and content morphemes in Tagalog-English do not appear to differ between the two modes of interaction. Once again, most of the contribution to the Chi-Square comes from the differences between observed and expected values for ELIs and inserts, where ELIs in Tagalog-English ICS are more likely to occur in the SMS mode and less likely to occur in the F2F mode. Inserts in Tagalog-English are more likely to occur in the F2F mode but less likely in the SMS mode.

6. Discussion

The first research question asked whether the types of morphemes in ICS and shifts in ML found in F2F conversations differed from those present in SMS exchanges. As shown in the Results Section, there were indeed significant differences between the observed and expected frequencies for these two modes of interaction. Most of the contribution to the fairly large Chi-Square in both the Spanish-English and Tagalog-English corpora is attributed to the differences in the types of ELIs, which are more likely to take place in the SMS mode than in the F2F mode. ELIs in ICS in Spanish-English often include the early system morphemes of the embedded language as in "Tengo quearreglar MY COUSIN'S COMPUTER," but there are also instances of ICS where the system morpheme of the ELI comes from the ML, such as "toy [estoy] seguraque Tina WOULD BE A PRETTY GIRL."

In addition to ELIs, inserts contribute to the fairly large Chi-Square in both the Spanish-English and Tagalog-English observed and expected frequencies in F2F interaction. During ICS episodes, inserts occur significantly more often in the F2F mode because lemmas are being rapidly activated in the BML of individuals as they communicate meaning in oral performance.

Also, in F2F interaction, processing time is constrained by the immediate nature of communication in which speech is characterized by false starts, disfluencies, hesitations, and frequent use of inserts generally embedded within the grammatical frame of a Matrix language. F2F interact ants do not have sufficient time to modify their utterances, so easily retrieved inserts are frequently used in conversations between bilinguals. These are used for emphasis, comic effect, to disparage someone (expletives), to keep the conversation going, or to ensure that someone has understood an utterance's intended meaning. No CS characterizes the output of older individual's aged60 and above as they interact with members of the second generation. They use Spanish almost exclusively in their interaction with younger family members and neighbors.

Members of different generations use different CS strategies in F2F interaction. This is also the case with CS in SMS conversations between and among members of different generations in the Tagalog-English corpus. A parent in his late 50s or 60s will use more Tagalog than English when texting a son or daughter, and these CS strategies yield different morphosyntactic patterns, which are manifested in Tagalog-English SMS conversations. As our analysis of all four corpora show, MLs shift frequently. In samples examined from the Miami corpus, English provides the sentential frame more often than Spanish does and this is also the case in the examples analyzed in the Spanish-English SMS mode. More ELIs is present in the SMS mode. As predicted by the BLA model, their presence in ICS can be explained by more time for cognitive processing and accessing lemmas while composing a message in response to a topic of conversation with a specific interlocutor.

In order to convey the intended meaning in an SMS conversation, texters often resorted to ELIs, as in the following example: "Ese para trabajando ALLTHE TIME" (in response to a question about an acquaintance, MC states that that guy is working all the time).In the SMS mode, there is more processing time so that semantic-pragmatic feature bundles at the conceptual level activate language-specific lemmas in the BML, where language-specific lemma activation takes place, mediating between conceptualization and text message production in this case(Wei,2006).When bilinguals engage in ICS via texting, they consciously or unconsciously choose one of the languages as the ML, which controls morph syntactic procedures, and theother one as the EL.

In Tagalog-English, also known as Taglish, the VSO order of Tagalog may be a factor that accounts for ICS and the frequency of ELIs present in the texting mode:"OH GOOD, WEARkarinnang COSTUME" (oh good,you wear a costume too).The team coding and tagging the Tagalog-English corpora in ICS episodes noticed that exact translations between the two languages would be largely incomplete without resorting to EL lexemes in orderto convey the intended meaning. According to Grosjean (1998, 2001/2007) when bilinguals interact and a particular word cannot be accessed in one language, CS takes place when that word is more readily available in the other language. He also notes that such decisions are deliberate. This appears to be the case with the usage of the phrase "sanduinkita" in a text message, which replaces a whole constituent, "Iwill pick you up."The message is as succinct as possible when using a simple phrase or a single word. Likewise, when switching to Tagalog as in "hatidkita," which means, "I will accompany you," the meaning is instantly conveyed to the person receiving the text message without havingto spells out an intended course of action in English. The second research question addressed the sociolinguistic functions found in these two different modes of communication. These were inferred by first identifying topics of conversation and undertaking a sequential analysis of code switched conversations in F2F and SMS interaction. Functions include various speech acts that allow bilinguals to argue, assert an opinion, clarify information for older individuals, coordinate social activities, criticize, disseminate and filter information, hold the floor or continue texting, negotiate meaning comprehension, and discursively maintain neighborhood ties and SMS social networks.

7. Conclusion

Generalization of these findings is limited to similar contexts and ethno-linguistic groups or members of similar communities of practice (e.g., computer science aficionados) given the geographically and culturally contextualized nature of both F2F and SMS conversations. The size of the SMS corpora also constrains the generalizability of these findings. Neither style shifting nor other individual features of SMS communication, such as the use of number-letter homophones or graphic icons were systematically investigated in this study. Using categories that are believed to be specific to the SMS medium of communication may yield different results for both form and function of CS when compared to CS in F2F interaction (Androutsopoulos, 2015; Blommaert, 2010; Morel et al., 2012.)

There are several implications for future studies of CS based on the findings of this study that use a combination of models and approaches to data analysis in two different modes of communication. First, the present study provides specific structural and lexical evidence to test the reliability of formal models traditionally used in investigations of CS inF2F contexts. Thus, in addition to the MLF model, Myers-Scotton& Jake's (2000, 2009) 4-M model of morpheme classification can be used to identify and classify different types of morphemes found in these two different modes of communication. However, there are limitations when applied to code switched SMS communication because it is difficult to classify initialisms, number-letter homophones, globalized formulaic expressions, icons and emoticons. Furthermore, consistent and dependable judgments among raters are difficult to obtain because of the model's complexity, which consists of high inference categories, and the significant variation that ICS data consistently exhibit (Chan,2009).

Statistically significant differences obtained when comparing two modes of communication. In the F2F mode, inserts predominate because speakers insert one or two words from one language(English) into the sentential frame of another (Spanish): JES: y y Paul esmuymuy HOTHEADED?They also use tags, interjections, or expletives in Spanish in utterance-initial position before an English clause or sentence (e.g.,MAR:coño YOU'LL THINK SHEHAS JUST ARRIVED HERE; or<Insert>coño</Insert> I just went blank).Thus, it seems that inserts are used to make a point, keep the conversation going, or hold the floor during memory lapses in a lengthy exchange. In the SMS mode, ELIs as compromise strategies occur more frequently when there is "sufficient congruence between the lexical-conceptual structures across the languages involved in ICS, but the predicate-argument structures differ"(Wei, 2006: 174).ELIs thus contributed to the high Chi-square values between the observed and expected frequencies found in the SMS mode but not in the F2F mode.

Second, the different frequencies of ELIs and inserts in these two modes of interaction can be explained as the result of lemma activation processes proposed by Wei's(2006, 2009) BLA model, which facilitates the interpretation of bilingual speech behavior and switches between languages as cognitively based strategies at an abstract level. This means that in order to convey their communicative intention, texters or speakers select specific semantic-pragmatic feature bundles pre-verbally at the conceptual level so as to use the word properly at the functional level in the syntactic environment created by the ML.

Third, concerning the sociolinguistic functions of both inter- and intrasentential CS, these were inferred by examining topics and sequences of code switched F2F and SMS conversations. For example, when discussing sexual or taboo topics in SMS exchanges, Spanish-English bilinguals switched to English: "Cada vez que voya Ecuador y se toca un TOPIC sexual I CAN ONLY GIVE MY OPINION IN ENGLISH porque me da cosa SAYING IT IN SPANISH LOL." (Every time I visit Ecuador and a sexual topic is raised, I can only give my opinion in English because I feel funny saying it in Spanish LOL.) Specific sociolinguistic functions of CS which were inferred from the topics of conversation among Spanish-English bilinguals in F2F interaction included arguing, asserting, criticizing, giving advice, gossiping, monopolizing the conversation, and planning. The sociolinguistic functions of Tagalog-English bilinguals engaged in F2F interaction included arguing about politics, asserting opinions, coordinating social activities, and discussing culture and food. On the other hand, more varied sociolinguistic functions were realized by Tagalog-English bilinguals who code switched via SMS as inferred from the data and results of an online survey. These included conveying important information from third parties and accommodating the linguistic needs of older parents.

Fourth, an important implication of the functions of CS in this study is that alternating languages extends the communicative competence of speakers and texters. When switching languages in ICS episodes, bilinguals convey information succinctly that would otherwise take tedious paraphrasing or circumlocutions, which is the case with various examples from all four corpora. This study contributes to extant literature on CS in two modes of communication. A close analysis of topics of code switched F2F and SMS conversations allowed us to infer specific sociolinguistic functions. CS is influenced by the interlocutor, the context, the topic under discussion, and the psychological state of texters or F2F interactants. The functions inferred ranged from those of a transactional nature (e.g., obtaining specific information from health practitioners or negotiating with third parties) to those of a strictly relational nature, such as arguing, criticizing, joking, requesting favors, rescheduling appointments, or planning social events. Future studies of CS in SMS or CMC should keep in mind recent arguments presented by scholars working with multilingual corpora that the diversity of grammatical forms and semiotic practices or hybridization of language present in digital communication calls for a rethinking or revision of traditional formal and functional categories of CS.

Notes: I gratefully acknowledge Dr. Yong Wang's assistance with SPSS and Dr. Longxing Wei's careful reading of the theoretical framework of this manuscript and examples provided.

References

Androutsopoulos, J. (2007).Language Choice and Code Switching in German-Based Diasporic Web Forums.In B. Danet & S. Herring (Eds.). The Multilingual Internet. New York: Oxford University Press, 340-361.

Androutsopoulos, J. (2015).   Networked multilingualism: Some language practices on Facebook and their implications. International Journal of Bilingualism, 19(2), 185-205.

Anis, J. (2007). Neography. In B. Danet& S. Herring (Eds.), The Multilingual Internet (pp. 87-115). New York: Oxford University Press.

Auer, P. (1995). The pragmatics of code-switching: A sequential approach. In L. Milroy & P. Muysken (Eds.), One speaker, two languages: cross-disciplinary perspectives on code-switching (pp.115-135). Cambridge: Cambridge University Press.

Auer, P. (1998). Introduction. Bilingual Conversation revisited. In P. Auer (Ed.), Code-switching in conversation (pp. 1-24). London: Routledge.

Biber, D., Conrad, S., &Reppen, R. (1998). Corpus Linguistics: Investigating language structure and use. Cambridge, UK: Cambridge University Press.

Blommaert, J.M.E. (2010). The sociolingustics of globalization. Cambridge: Cambridge University Press.

Bautista, M.L.S., (1990). Tagalog-English code-switching revisited. Philippine Journal of Linguistics, 21(2), 15-29.

Bautista, M. L.S. (2004). Tagalog-English code switching as a mode of discourse. Asia Pacific Education Review, 5, (2), 226-233.

Chan, B. H. (2009). Code-switching between typologically distinct languages. In B. E. Bullock & A. J. Toribio (Eds.), The Cambridge Handbook of Linguistic Code-switching. New York: Cambridge University Press, 182-198.

Chiluwa, I. (2008).   Assessing the Nigerianness of SMS text-messages in English. English Today 24(01):51 - 56 ·

Danet, B., & Herring, S. (2007). Introduction. In B. Danet & S. Herring (Eds.), The Multilingual Internet(pp. 3-39). New York: Oxford University Press.

Deumert, A & Masinyana, S.O. (2008). Mobile language choices - The use of English andisi Xhosa in text messages (SMS): Evidence from a bilingual South Africansample. *English World-Wide*, 29(2), 117-147.

Dürscheid, C. & Stark, E. (2011). SMS4sciencecorpus-Based Texting Project and the Specific Challenges for Multilingual Switzerland. In C. Thurlow, K. Mroczek (eds.) Digital Discourse. Language in the New Media, (pp.299-320). Oxford: Oxford University Press.

Fairon, C.,Klein, J.R., & Paumier, S (2006).Le langagesms. Étude d'un corpus informatisé à partir de l'enquête "Faites don de vossms à la science."Louvain-la-Neuve, Presses universitaire de Louvain.

Gafaranga, J. (2009). The Conversation Analytic model of code-switching.   In B. E. Bullock & A. J. Toribio (Eds.), The Cambridge Handbook of Linguistic Code-switching (114-126). New York: Cambridge University Press.

Gumperz, J.J (1982). Discourse strategies. Cambridge, UK: Cambridge University Press. Gumperz, J. J. (2001). Interactional sociolinguistics. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), The Handbook of Discourse Analysis (pp. 215-228). Oxford: Blackwell.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. Bilingualism: Language and Cognition, 1(2), 131-149.

Grosjean, F. (2001/2007). The bilingual's language modes. In J. Nicol (Ed.), One Mind, two languages: Bilingual language processing (1-22). Oxford: Blackwell. Reprinted in Li Wei (Ed.), The Bilingualism Reader (pp. 428-49). London and New York: Routledge.

Marzuki, E. & Walter, C. (2013).English and Malay Text Messages and What They Say about Texts and Cultures.Open Journal of Modern Linguistics3(4), 295-304.

Milroy, L. (1987). Language and Social Networks. Oxford: Blackwell.

Morel, E., Bucher, C., Doehler, S. P., & Siebenhaar, B. (2012).   'SMS communication as plurilingual communication: hybrid language use as a challenge for classical code-switching categories'. Linguisticae Investigationes 35(2), 260-288.

Muysken, P. (2000). Bilingual speech: A typology of code-mixing. Cambridge: Cambridge University Press.

Myers-Scotton, Carol (1993).Common and Uncommon Ground: Social and Structural Factors in Codes witching. *Language in Society*, 22, (4), 475-503.

Myers-Scotton, C., & Jake, J. (2000).Testing the 4-M model: An introduction. *The* International Journal of Bilingualism, 4(1), 1-8.

Myers-Scotton, C., & Jake, J. (2009). A universal model of code-switching and bilingual language processing and production. In B. E. Bullock& A. J. Toribio(Eds.), The Cambridge Handbook of Linguistic Code-switching. New York: Cambridge University Press, 336-357.

Paolillo, J.C. (2007).In B. Danet & S. Herring (Eds.), The Multilingual Internet (408-430). NewYork: Oxford University Press.

Poplack, S. (1980).Sometimes I'll start a sentence in Spanishy terminoenespañol: toward a typology of code-switching. Linguistics, 18, 581-618.

Poplack, S. (1988). Contrasting patterns of code-switching in two communities. InM. Heller(Ed.), Codes witching: Anthropological and sociolinguistic perspectives. Berlinand New York: Moutonde Gruyter, 215-244.

Rafael, V.L. (1995). Taglish, or the phantom power of the lingua franca. Public Culture, 8(1), 101-126.

Simango, S.R. (2011). When English meets isiXhosa in the clause: An exploration into the grammar of code-switching. Southern African Linguistics and Applied Language Studies 29 (2), 127-134.

Smedley, F.P. (2006). Code-switching and Identity on the Blogs: An Analysis of Taglish in Computer Mediated Communication (Unpublished Master's dissertation). Auckland University of Technology.

Wei, L. (2006). Intrasentential Codeswitching as Conceptual Projection of Lemmas in the Bilingual Mental Lexicon. Journal of Cognitive Science, 7, (2), pp.149-178.

Wei, L. (2009). Code-switching and the bilingual mental lexicon. In B. E. Bullock & A. J. Toribio (Eds.), The Cambridge Handbook of Linguistic Code-switching (pp. 270-88). New York: Cambridge University Press.

Woolard, Kathryn (2004). "Codeswitching". In A. Duranti. A Companion to Linguistic Anthropology (pp.73-94). Mass.: Blackwell.

Zentella, A. C. (1982). Code-switching and interaction among Puerto Rican children. In J. Amastae& L. Elias-Olivares (Eds.), Spanish in the

United States: Sociolinguistic aspects (pp. 354–385). Cambridge: Cambridge University Press. Zentella, A.C. (1997). Growing Up Bilingual. Malden, MA: Blackwell