Sometimes Ngo Zau Start to Gong Chinese

The role of pragmatics and information structure in the syntax of codemixing

Andrew Peters*– abpeters@yorku.ca

WISSLR, Mar. 5th, 2016

1 Introduction

1.1 Background on Codemixing Research

Research into bilingual syntax and intra-sentential codeswitching (or codemixing) has focused around three major theoretical avenues to approach regularities in the behaviour:

1. Approaches that posit meta-structures and domain-specific mechanisms (eg. Poplack, 1980, 2004; Myers-Scotton, 1993, 1995 *inter alia*)
2. Approaches that posit no domain-specific mechanisms, and argue that only Syntax itself constrains codemixing (eg. MacSwan, 1999, 2009; Chan, 1998 *inter alia*)
3. Discourse-affective approaches such as of Gumperz (1972 *inter alia*), Auer (1988, 1995 *inter alia*) and Poplack (1980, 1998, 2004 *inter alia*) who posit that discourse plays the greatest role in informing the structure

I take second approach, the generative approach, as theoretically preferable and a starting point of investigation.

Example of a Generative approach to Codemixing:

- MacSwan argues that constraints on codemixing result from the same mechanisms of interface at PF that check or value features at spell out.
- For example, English pronouns undergoe head-movement to T, which causes a crash at PF as heads are inputs to PF (Chomsky, 1995), and must be of a single language.
- However, in Cantonese no agreement morphology exists. Huang (1984) and Chou (2013) analyse Chinese as a language which does not inherit phi-features on T.

(1)[1] is grammatical in Cantonese-English Codemixing:

(1) li1 go3 leoi5 zai2 le1, **keoi5 pass** zo2 ge3
This CLASS. girl TOP.PRT., **she pass** PFV.ASP. POSS.
"This girl, she passed"

1.2 On Cantonese, and Data for the Present Study

I will examine mixing behaviour from Cantonese-English bilinguals. While both are fairly isolating languages with relatively little inflectional morphology, Cantonese is much more so than English (Matthews & Yip, 1994).

Cantonese is:

- a dialect of the Yue branch of Chinese spoken in southern China
- highly discourse-configurational(Matthews & Yip, 1994[2]; Huang, Li, & Li, 2009): Chinese languages were one of the first examined by linguists like Li and Thompson (1976) who forwarded the topic-comment model

I examine data from a corpus of Heritage Cantonese speakers from Toronto, from the Heritage Language Variation and Change project (Nagy, 2011). The interviews chosen for investigation are from second and third generation speakers of heritage-Cantonese, born and raised in Toronto, who are natively-bilingual in English and Cantonese. In surveying data from five speakers in the corpus, nearly 400 instances of intra-sentential codeswitching (codemixing) were catalogued.

1.3 Cantonese-English Codemixing

Cantonese-English Codemixing permits switching of pronominal subjects:

(2) daan6 hai6 **o5 dei6** sing4 jat6 **complain** le3
But **we** every day **complain** PART.
"But we complain every day."

*Thanks for comments and critiques from: Dr. Angermeyer, Gavin Bembridge, and Holman Tse. Thanks to Dr. Nagy for the opportunity to work with the HLVC project and access to the corpus.

[1] All Cantonese speech has been transcribed using the Jyutping romanisation paradigm (LSHK, 2002).

[2] see (Kiss, 1994); (Lambrecht, 1994) for more information

Cantonese-English codemixing also permits subject pro-drop:

(3) Gan1 zyu6 Ø **start to** gong2 o5 dei6 go3 relationship le1
Then Ø **start to** say we CLASS. **relationship** PRT.
"Then [we] start to talk about our relationship."

Notice as well that the matrix verb "start" selects and English infinitival TP, and that the language of the T head may differ from that of the embedded verb.

There usually is not exhibited inflectional morphology on English verbs, and past tense morphology never appears even when the event occurred in the past and has Cantonese aspectual markers:

(4) gei2 lin4 zau6 **give up** zo2 MSN laa1.
Several year just **give up** PFV.ASP. MSN EMPH.PRT.
"[I] gave up MSN for several years."

(5) **Settle** zo2 zau6 bun1 heoi3 Saskatchewan hoi1 jat1 go3 farm.
Settle PFV.ASP. just move go Saskatchewan open one CLASS. farm
"[They] settled and moved to Saskatchewan to open a farm."

Not only does Cantonese-English codemixing permit pro-drop of subjects, but it permits object-drop as well, even when the main verb is English:

(6) jyu4 go2 jau5 gei1 wui6 dou1 wui5 **send** Ø ge2.
If have opportunity also will **send** Ø PRT.
"If [I] had the opportunity, I would send [them]."

This last case is in line with Cantonese-style discourse-motivated argument drop (Matthews & Yip, 1994).

When English nouns are mixed into Cantonese syntax, they may appear either with or without plural morphology where it would be obligatorily called for in monolingual syntax.

(7) go2 di1 **neighbour** lo3, like bong1 each other
DEM. CLASS. **neighbour** PRT. like help each other
"Those neighbours, like, help each each other."

(8) wui6 jau5 hou2 do1 go2 di1 **thoughts**
will have very many DEM. CLASS. **thoughts**
"[They] will have a lot of those thoughts"

2 The Problem: Nouns in Bilingual Syntax

2.1 The Cantonese DP

In Cantonese, the DP consists of a Demonstrative, a Numeral(NUM), a Classifier (CL) and a Noun (Matthews & Yip, 1994), an example of which is provided in (9).

(9) go2 saam1 zoeng1 toi4
DEM. three CL. table
"Those three tables"

- There exist many classifiers of different types, but a general classifier *go3* is the most common.
- There is a non-singular general classifier, *di1* which gives a non-singular reference to the noun.
- We will return to the nature of this quantity feature in section (2.3)

2.2 Overview of the Data

When English nouns are mixed into Cantonese-English speech, plural morphology is optional where it would otherwise be obligatory in monolingual English speech.

(10) go2 di1 **rule** le1, keoi5, so we had, they had to, they had to hide
DEM. CLASS. **rule** PRT., he- so we had, they had to, they had to hide
"[There were] those rules, they- so we had, they had to, they had to hide"
(definite, count noun, plural, no plural morphology)

(11) o5 zung1ji3 tai2 go2 di1 **electronics** lo1
I like look DEM. CLASS. **electronics** PRT.
"I like to look at the electronics"
(definite, mass noun, plural, with plural morphology)

(12) o5dei6 teng1 dou2 o5dei6 go2 di1 **family friend** le1
We listen PRT.achieve we DEM. CLASS. **family friend** PRT.
"We heard our family friends."
(definite, count noun, plural, no plural morphology)

(13) jau5 di1 **case** lo1
have CLASS. **case** PRT.
"There are some cases."
(indefinite, count noun, plural, no plural morphology)

(14) lei5 jiu3 faan1 gung1 gon2 si4 jiu3 jung6 lo1. Ze1
You need return work CONTR: DEM. TIME time need use PRT.. That
hai6 di1 **skills.**
is CLASS. **skills.**
"What you need when working. That is, those skills."
(indefinite, count noun, plural, with plural morphology)

2.3 Previous Studies: Chan's (1998) Analysis

Chan (1998) posed a solution to the optionality of plural morphology:

- Cantonese D heads do not project a [PLURAL] feature
- Cantonese classifiers and quantifiers project a [QUANTITY] feature, which is not the same as a plural feature, but is correlated with similar semantic content
- In codemixing, English nouns that project a [+PLURAL] feature or are underspecified for plurality may be selected

Some problems exist.

- Chan's view of the select function is somewhat simplistic or outdated. In most modern syntactic analyses that assume Distributed Morphology (DM), the Late Insertion Hypothesis holds.

(15) **The Late Insertion Hypothesis**
"... syntactic categories are purely abstract, having no phonological content. Only after syntax are phonological expressions, called Vocabulary Items, inserted in a process called Spell-Out. ... " (Harley & Noyer, 2003)

- Chan (1998) inherits problem of how to attribute the linguistic elements in the syntax to any particular language.
- Resolved under DM: terminal nodes are abstract objects which project features.

However: In the data surveyed from the HLVC corpus, 44 instances of referentially plural English nouns were found, 33 of which exhibited no plural morphology, 11 with plural morphology.

The question now becomes not "why are some nouns underspecified for [+plural]?" but rather "**what licenses plurality when it *is* specified for?**"

3 Solving the Problem

3.1 Review: The structure of Cantonese DP's; Questions to Answer

- Cantonese DP's: a Demonstrative, a Numeral(NUM), a Classifier (CL) and a Noun
- [QUANTITY] feature of some type may appear on Cantonese Classifiers
- Cantonese nouns do not value number-ϕ-features at spell-out.
- English DP's have a full set of [PLURAL] and ϕ-features which value into plural morphology

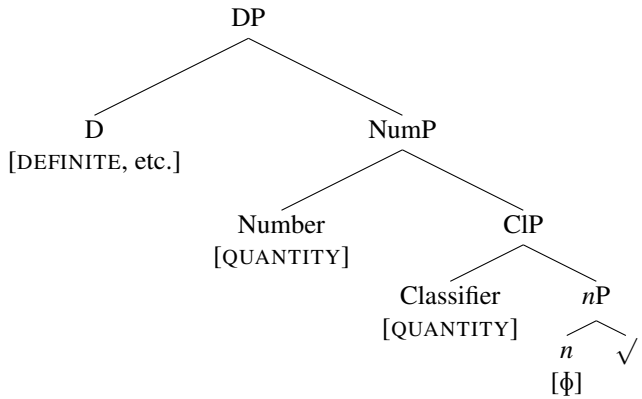

Figure 1: Cantonese-Style DP

Positing this structure, our questions about English Nouns and their plural morphology in the bilingual syntax are now:

1. **Structurally**: How do English [PLURAL] features get valued in bilingual syntax to produce plural morphology?
2. **Pragmatically**: What licenses the seemingly uncommon selection of [+PLURAL] featural projections? Or: In what contexts does plural morphology appear, and how is it accommodated by speakers?
3. **Discursively**: How does this relate to the observation by researchers such as Gumperz (1972 *inter alia*), Auer (1988, 1995 *inter alia*) and Poplack

(1980, 1998, 2004 *inter alia*) who hold that the discursive function of the code-switch is the greatest driving factor in determining its realisation?

3.2 The Structural Question

Chan (1998) may well have been right about the [QUANTITY] feature valuing the [PLURAL]. However, as seen in figure (1), ϕ-features project on *n*, which takes the nominal root as its complement.

The story of an English noun **without plural morphology** looks identical to that of a Cantonese noun in monolingual Cantonese Syntax: a nominal root is selected as complement of *n* which projects the minimal set of ϕ-features (sans [PLURAL]), and when the derivation reaches spell out, the speaker selects an English noun from the Encyclopaedia. The nominal root combines with the *n* head, and no plural morphology manifests.

The story of an English noun **with plural morphology**, however:

1. A nominal root is selected as complement of *n* which **does** project a [+PLURAL](for some reason?).
2. The derivation reaches spell-out and a noun is mapped to the root from the English Encyclopaedia.
3. The [+QUANTITY] feature checks and values the [+PLURAL] feature, producing English plural morphology on the English noun.

A question remains: **Where, and why is the [+PLURAL] feature projected?**

3.3 The Pragmatic Question

Key Observation: **plural morphology only appeared on focussed elements, and never on topics.**

- Féry & Krifka (2008) provide a definition of both Topic and Focus in terms of the Common Ground (CG)
- CG: originally introduced by Stalnaker (1974) to describe the set of information that has been introduced into the discourse and is available to the interlocutors

(16) **Definition of Topic**
"The topic constituent identifies the entity or set of entities under which the information expressed in the comment constituent should be stored in the CG content." (Fery & Krifka, 2008)

Kiss (1994) identifies at least two types of focus in discourse-configurational languages: new-information (Rhematic) focus, and contrastive (Identificational) focus.

Both types of focus license projection of [+PLURAL] features.

(17) go2 di1 **rule** le1, keoi5, so we had, they had to, they had to hide
DEM. CLASS. **rule** PRT., he- so we had, they had to, they had to hide
"[There were] those rules, they- so we had, they had to, they had to hide"
(Topicalised DP, no plural morphology)

(18) go2 di1 **neighbour** lo3, like bong1 each other
DEM. CLASS. **neighbour** PRT. like help each other
"Those neighbours, like, help each each other."
(Topicalised DP, no plural morphology)

(19) Hai6 **clowns** hou2 geng1 lo1.
be **clowns** very scary EMPH.-PRT.
"It's clowns that are scary."
(Identificational Focus, with plural morphology)

(20) Ah, jan1 wai6 keoi5 dei6 ge3 **accents**
Ah because they POSS.-PRT. **accents**
"Ah, because of their accents"
(Identificational Focus, with plural morphology)

(21) o5 zung1ji3 tai2 go2 di1 **electronics** lo1
I like look DEM. CLASS. **electronics** PRT.
"I like to look at the electronics"
(New Information / Rhematic Focus, with plural morphology)

The behaviour seems to be focus-sensitive. That is, when the information exists already in the CG, that is, of a Topic, plurality is never specified for. However, when the information does not exist in the CG, and the noun is referentially plural, a [+PLURAL] feature is projected.

3.4 The Discourse Question

Once the influence of pragmatics and information structure has been recognised as so influential in the codeswitching behaviour, it becomes difficult to turn away from the comments of researchers like Gumperz, Auer and Poplack who hold that the discursive function is of importance in determining the form of the switch.

If the information structure, pragmatics and context-sensitive semantics can be shown to play an equally important role in other aspects of bilingual syntax, the long-standing observation by Gumperz, Poplack, Auer and many others that code-switching may be more or less discourse-affective will turn out to be a symptom of the syntax pragmatics interface. It is no surprise that contrastive and new-information focussed elements will be symptomatically be associated with the types of switches that Poplack and Gumperz take to be highly discourse-affective or "flagged" (Poplack, 1980) switches.

That the unavailability of information in the common ground conditions the optional expression of plural morphology speaks to the influence of the discursive function on the information structure of a sentence. Where speakers make use of this influence for purely discursive ends, it seems they are exploiting the symptoms of the syntax-pragmatics interface.

References

Auer, P. (1988). A conversation analytic approach to code-switching and transfer. In M. Heller (Ed.), *Codeswitching* (pp. 187–214). Berlin,: Mouton de Gruyter.

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.

Chan, B. H.-S. (1998). Functional heads, cantonese phrase structure and cantonese-english code-switching. *UCL Working Papers in Linguistics*, *10*.

Chomsky, N. (1995). *The minimalist program*. Cambrdige, MA: MIT Press.

Chou, C.-T. (2013). Phi-agree, a-movement and complementizer-tense relations in chinese. *Ph.D. Dissertation*.

Fery, C., & Krifka, M. (2008). Information structure: Notational distinctions, ways of expression. In P. van Sterkenburg (Ed.), *Unity and diversity of languages* (p. 123-136). Amsterdam: John Benjamins.

Gumperz, J., & Blom, J.-P. (1972). Social meaning in linguistic structure: code-switching in norway. In J. Grumperz & D. Hymes (Eds.), *Directions in sociolinguistics* (pp. 407–434). New York,: Holt, Rinehart and Winston.

Harley, H., & Noyer, R. (2003). Distributed morphology. In L. Cheng & R. Sybesma (Eds.), *The second glot international state-of-the-article book* (p. 463-496). De Gruyter Mouton.

Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, *15*(4), 531–574.

Huang, C.-T. J., Li, Y.-H. A., & Li, Y. (2009). *The syntax of chinese*. Cambridge, UK: Cambridge University Press.

Kiss, K. E. (1994). *Discourse configurational languages*. Oxford University Press.

Lambrecht, K. (1994). *Information structure and sentence form*. Cambridge, UK: Cambridge University Press.

Li, C. N., & Thompson, S. A. (1976). Subject and topic: A new typology of language. In *Subject and topic: [papers]*. New York,: Academic Press.

LSHK. (2002). *Guide to lshk cantonese romanization of chinese characters* (2nd ed.). Linguistic Society of Hong Kong.

MacSwan, J. (1999). *A minimalist approach to intrasentential code switching: Spanish-nahuatl bilingualism in central mexico*. London & New York,: Routledge.

MacSwan, J. (2009). Generative approaches to code-switching. In B. E. Bullock & A. J. Toribio (Eds.), *The cambridge handbook of linguistic code-switching*

(pp. 309–335). Cambridge,: Cambridge University Press.
Matthews, S., & Yip, V. (1994). *Cantonese: A comprehensive grammar*. London,: Routledge.
Myers-Scotton, C. (1993). *Duelling languages: Grammatical structures in codeswitching*. Oxford,: Oxford University Press.
Myers-Scotton, C. (1995). A lexically-based production model in code-switching. In L. Milroy & P. Muysken (Eds.), *One speaker: Two languages* (pp. 233–256). Cambridge,: Cambridge University Press.
Nagy, N. (2011). A multilingual corpus to explore variation in language contact situations. *Rassegna Italiana Di Linguistica Applicata*, *43*, 65-84.
Poplack, S. (1980). Sometimes i'll start a sentence in spanish y termino en espanol. *Linguistics*, 581–618.
Poplack, S. (2004). Code-switching. In U. Ammon, N. Dittmar, K. Mattheier, & P. Trudhill (Eds.), *Soziolinguistik an international handbook of the science of language.* Berlin,: Walter de Gruyter.
Poplack, S., & Meechan, M. (1998). Introduction: How languages fit together in codemixing. *International Journal of Bilingualism*, *2*, 127–138.
Stalnaker, R. (1974). Pragmatic presuppositions. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and philosophy* (p. 197-214). New York: New York University Press.