Psycholinguistic and Corpus Approaches to Codeswitching

Melinda Fricke, University of Pittsburgh
melinda.fricke@pitt.edu
Course Goals

(1) Start with a brief overview of the vast psycholinguistic literature on bilingual language use (mostly production, a tiny bit of comprehension).

(2) Then dip our toes into the observational, primarily linguistic literature on codeswitching (CS).

(3) Finish by looking at some attempts to integrate naturalistic observation of CS phenomena with psycholinguistic models.

(4) End up with a preliminary project proposal for a study that would help to advance the current psycholinguistic understanding of CS.
Course Format

• In general:
  • Cover 3-4 papers per day (10 min presentation + 15-25 min discussion).
  • You are not expected to do all of the readings, but it will be helpful if you can skim them!
  • 19 papers / 9 students = ~ 2 paper presentations per student.

• Today:
  • Brief overview of course topics and papers so we can decide who’s presenting what.
  • I will present the Costa (2009) and Meuter & Allport (1999) readings to give you an idea of what you should cover in your presentations.
Schedule of Topics

• Days 1 – 3: Psycholinguistics of Bilingualism

7/6  **Costa (2009):** Review article from the *Handbook of Bilingualism: Psycholinguistic Approaches* on psycholinguistic models of bilingual (lexical) production.

**Meuter & Allport (1999):** Classic (widely cited) study looking at “switch costs” in a cued language switching task.

7/10 **Green (1998):** Very widely cited model of bilingual language control.

**Costa & Santesteban (2004):** A more recent look at switch costs, again using a cued language switching task.

**Goldrick et al. (2014):** Yet another look at switch costs, this time measuring phonetic production (voice onset time) instead of response time.
Schedule of Topics

• Days 1 – 3: Psycholinguistics of bilingualism

7/13  

**Grosjean (2001):** An overview of Grosjean’s (now classic) idea of “language modes”.

**Loebell & Bock (2003):** A first look at syntactic priming across languages.

**Salamoura & Williams (2007):** More syntax: evidence for cross-language activation from cognates, grammatical gender, and adjective agreement.

**Hartsuiker & Pickering (2008):** A nice (short) overview of models of bilingual syntactic representation.
Schedule of Topics

• Days 4 – 5: Corpus-based examinations of codeswitching phenomena

7/17

Li (2013): Review chapter from the Oxford Handbook of Sociolinguistics on CS (watch out for typos, unfortunately).

Lanza (1992): Classic paper on whether language switching in very young children (2-year-olds) should be considered CS.

Poplack (1980): The classic CS paper, exploring various aspects of systematicity in English-Spanish CS among Puerto Ricans in NYC.

Angermeyer (2002): Argues for an important role for the discourse functions of CS, using data from a German/French/English-speaking family.
Schedule of Topics

- Days 4 – 5: Corpus-based examinations of codeswitching phenomena

  7/20

  **Stammers & Deuchar (2012):** Argues against Poplack’s Nonce Borrowing Hypothesis using data from Welsh-English CS.

  **Poplack, Zentz, & Dion (2012):** Argues against the idea that CS necessarily leads to language convergence using data on preposition stranding in Québeçois French.

  **Myslín & Levy (2015):** Argues (from an information-theoretic framework) that CS is used to highlight “new” or “more informative” information, using data from Czech-English CS (and a cool experiment).
Schedule of Topics

• Days 6 – 7: Psycholinguistically oriented work on CS


N.B. I will not be here this day!

The remainder of class time will be devoted to group work on your project proposals.
Schedule of Topics

• Days 6 – 7: Psycholinguistically oriented work on CS

7/27

Guzzardo Tamargo et al. (2016): Combination corpus study and eye tracking study, looking at how statistical regularities in CS impact reading times.

Broersma & de Bot (2006): Corpus study providing the first quantitative evidence for “triggered” CS.

Li (1996): Early study looking at the role of phonetic integration in the recognition of “CS” speech.

Kleinman & Gollan (2016): A further update on switch costs in the language switching paradigm (with no cues!).
Schedule of Topics

• Day 8: Discussion of project proposals

7/31 No assigned readings! But see the folder labeled “Additional Papers of Interest” for inspiration.

Presentations of project proposals, with time for feedback and discussion.

Final write-ups due to me (melinda.fricke@pitt.edu) by midnight on August 1st. Absolute maximum of 1000 words!
**Review or Modeling Articles**

1. **Summary of topic/purpose.**
   What general topic does the review address, or what phenomenon is the model designed to account for?

2. **Summary of argumentation.**
   How does the paper make the problem more tractable? Briefly summarize the major subparts of the paper.

3. **Evaluation: What do you like about the paper?**
   These points could be organizational or theoretical.

4. **Evaluation: What do you dislike about the paper?**
   Are there any gaps, or arguments that just don’t make sense to you? How might you amend the proposal or conclusions?
Paper Presentation Format

Empirical Articles

1. **Summary of topic/purpose.**
   - **Big question:** Overarching theoretical motivation for the study.
   - **Little question:** What is the specific research question that will provide evidence *bearing on* the big question?

2. **Summary of methods and findings.**
   - How did the researchers implement their study? Who were the participants, and what were the main results?

3. **Evaluation: What do you *like* about the paper?**
   - These points could be organizational, theoretical, or methodological.

4. **Evaluation: What do you *dislike* about the paper?**
   - Are there any gaps, or arguments that just don’t make sense to you?
   - Any ideas for how you could implement the study differently to address the issues you see?
Costa (2009)

(1) Summary of topic/purpose.

• What are the representations and processes involved in bilingual language production (esp. of single words)?

• When bilinguals intend to speak in a “target language”, to what extent does the “non-target” language become active (and how can we tell)?
Costa (2009)

(2) Summary of argumentation.

• 3 levels of linguistic representation: semantic, lexical, and phonological.

• The time course of when these become active (“activation flow”) is a source of controversy.
  • Most psycholinguistic models of bilingualism agree that (lexical?) activation is “non-selective”: Representations in both languages become active even when the target language is known.
  • This means there needs to be a “selection mechanism” to explain why the target language is usually successfully produced.
Costa (2009)

A few words on “activation flow”...

- language specific vs. non-specific
- this question applies to lexical and to phonological activation
- discrete vs. cascaded flow (does selection act as a “filter” or not?)
Costa (2009)

A few words on “selection mechanisms”...
Some experimental evidence: from semantic to lexical

- Poulisse & Bongaerts (1994) analyzed slips of the tongue in 45 late Dutch-English bilinguals:
  - High proficient: 16 out of 3361 L1 intrusions while speaking L2 (= 0.5%)
  - Low proficient: 246 out of 2795 (= 9%)
  - Non-selective lexical activation for all bilinguals, but proficiency mitigates chance of selecting wrong language.
Costa (2009)

Some experimental evidence: from lexical to sublexical

• Lots of evidence for cognate effects in language production
  • Costa et al. (2000) and many others have found that cognates are named faster than noncognates.
  • Gollan & Acenas (2000) found that cognates are less susceptible to tip-of-the-tongue states.

• These effects are generally interpreted as meaning that activation flows from lexical representations, down to phonological representations, in a non-selective way (across both languages).
Costa (2009)

Some experimental evidence: is inhibition necessary?

• Meuter & Allport (1999) argued in favor of inhibitory control in proportion to the strength or dominance of the two languages: the more dominant L1 is, the more inhibition it requires, so the more difficult it is to switch back into L1.

• N.B. Costa seems to attribute some conclusions to M&A that I don’t think they draw (namely, the idea that inhibition is item-specific and reactive). He also doesn’t allow for the idea that bilinguals can be better or worse in inhibitory control.

• He proposes that a “checking mechanism” could do the work of inhibitory control, and he points out that inhibitory mechanisms are not as widely accepted in monolingual language production.
(3) Evaluation: What do you like about the paper?

• Starts with a general overview of the issues involved, then provides a more detailed summary of evidence at the end.

• This makes it more readable: you don’t have to wade through all of the specific evidence before you can grasp the more general concepts and arguments.

• Fairly balanced; you can’t necessarily tell which side Costa would come down on.
(4) Evaluation: What do you dislike about the paper?

• This may be my bias coming in, but... I think the idea of a language-specific selection mechanism is somewhat non-standard.
  
  • What exactly does this mean? How would this be implemented? Something to do with attention and the ability to use contextual information effectively? This isn’t laid out in great detail, making it more difficult to understand.

• Great introduction to general concepts, but seems to confuse some of the finer points (e.g. Meuter & Allport’s explanation of their results).
Costa (2009)

Questions? Comments? Other reactions?
Meuter & Allport (1999)

(1) Summary of topic/purpose.

Big question: Does bilingual language selection depend on more general mechanisms? (“similar in kind to those responsible for the control of task set in other monolingual and/or nonlanguage task domains”)

Little question: Is there an asymmetric switch cost for the more dominant language (L1) as compared to the less dominant (L2)?
Meuter & Allport (1999)

(2) Summary of methods and findings.

• 16 non-balanced bilinguals, fluent in English and an [Indo-Eur] L2 (French, German, Italian, Portuguese, Spanish).

• Main experiment: ~2000 trials of digit naming, language cued by background color.

• “Runs” of varying length, with varying number of language switches.
Meuter & Allport (1999)

The diagrams illustrate the relationship between response latency (ms) and the conditions of Non-Switch and Switch for different categories: COMBINED, FIRST, SECOND, and THIRD. Each category is represented by a graph showing the trend of increasing response latency as the Switch condition occurs, with data points labeled L1 and L2.
Meuter & Allport (1999)
Meuter & Allport (1999)

(2) Summary of methods and findings.

• Hypothesis confirmed: Language switching shows the same asymmetric switch cost as other non-linguistic tasks.

• Therefore language control seems to rely on “domain-general” mechanisms of task control.
Meuter & Allport (1999)

(3) Evaluation: What do you like about the paper?

• Clear explanation of a counter-intuitive hypothesis and finding.

• Relates language to cognitive processing more generally.

• Use of participants with different language backgrounds is an interesting way to “control” for some unwanted effects; increases confidence in generality of findings.
Meuter & Allport (1999)

(4) Evaluation: What do you dislike about the paper?

• Not a criticism of the paper, but rather a criticism of language switching studies: To what extent can we believe this has anything to do with the day-to-day experience of being bilingual?

• In other words: Does this paper have anything to do with CS?