Covariance of syntactic and phonological contact effects in Eastern Cham

Kenneth Baclawski Jr.
University of California, Berkeley

NWAV 45, November 6, 2016

Intro

- Two phonological variables covary with one syntactic variable
  - Categorical presence or absence of construction
  - In contrast with alternations typical of socio-syntax
- Eastern Cham contact with Vietnamese has phonological and syntactic exponents
- Further evidence for the study of syntactic variation (Grondelaers & Speelman 2007, a.o.)
- Generally supports a competing grammar approach (cf. Kroch 1994)

Outline

- Eastern Cham/Vietnamese bilingualism
- Phonological variables
  - /r/, /j/
  - /ŋ/
- Syntactic variation
  - Discourse anaphora
  - *Wh*-topicalization
- Covariance and socio-syntax

Eastern Cham & Vietnamese

- Eastern Cham
  - Austronesian: South-Central Vietnam
  - 100,000 speakers, all bilingual with Vietnamese
  - Endangered (UNESCO)
    (cf. Brunelle & Văn Hẳn 2015)
- Vietnamese
  - Austroasiatic
  - 100 million speakers worldwide
  - Dominant political/socioeconomic language of Vietnam
- Typological similarities due to thousands of years of contact (Thurgood 1999)

Eastern Cham diglossia

- Exhibits stable diglossia
  - Brunelle (2009) on one variable: monosyllabification
  - Baclawski (2016) reports two other phonological variables
  - Emergence in Aymonier & Cabaton (1906), stable by Blood (1961), Alieva (1991, 1994)
- Register #1: Formal speech with connections to Cham script
- Register #2: Colloquial speech with Vietnamese contact effects

Predictions

- Intense language contact can lead to morphosyntactic contact effects (e.g. Thomason & Kaufman 1988)

- In a unidirectional contact situation, contact effects are liable to have parallel distributions

Data collection

- 30 native Eastern Cham speakers
  - 15 male, 15 female
  - Aged 18-79 (median = 22)
  - Raised in Cham villages, Ninh Thuận Province
- Survey targeting phonological variables
- Syntactic elicitation

Phonological variables

- 1. /r/, /j/
- 2. /ŋ/ / V_{rd} __
- Both are Vietnamese contact effects (Baclawski 2016)
- Both exhibit significant inter- and intra-speaker variation

/r/, /j/ in Eastern Cham

- Blood (1961) reports variation of /r/, /j/ in onsets:
  - [r] ~ [j] ~ [z]
- Baclawski (2016) reports similar facts:
  - [r] ~ [rʒ] ~ [j] ~ [z] ~ [jʒ]

- /rːàʔ/ 'market'
  - [rːàʔ] ~ [rʒːàʔ] ~ [jːàʔ] ~ [zːàʔ] ~ [jʒːàʔ]
- /jːaw/ 'wood'
  - [rːaw] ~ [rʒːaw] ~ [jːaw] ~ [zːaw] ~ [jʒːaw]

/j/ in Vietnamese

- Tran & Norris (2010) report for /j/ in onsets:
  - [j] (Southern dialects) ~ [z] (Northern dialects) ~ [z̺] (observed in Central)
  - Baclawski (2016) conjectures that all these forms are present in South-Central Vietnam
    - Dialectology is urgently needed in Vietnam in general

- /jə̀ː/ <giờ> 'hour': [jə̀ː] ~ [zə̀ː] ~ [z̺ə̀ː]

/ŋ/ / V$_{rd}$ __ in Eastern Cham

- Blood (1961) reports variation of /ŋ/ / V$_{rd}$ __ :
  - [ŋ] ~ [ŋ͡m]
- Baclawski (2016) reports the same facts

- /pùŋ/ 'top.of'
  - [pùŋ] ~ [pùŋ͡m]

/ŋ/ / V$_{rd}$ __ in Vietnamese

- Allophony of /k,ŋ/ reported since the 1600's (Jacques 2002, citing the Portuguese missionary de Rhodes; Thompson 1965)
  - /k,ŋ/ → [k͡p, ŋ͡m] / V$_{rd}$ __

- /ɔ́k/ <óc> 'snail' → [ɔ́k͡p]
- /oŋ/ <ông> 'grandfather' → [oŋ͡m]

  - Note: Eastern Cham lacks final /k/

Summary

- For /r/,/j/: [j], [z], and [jʒ] are contact variants
  - [r] and [rʒ] are non-contact variants

- For /ŋ/: [ŋ͡m] is a contact variant
  - [ŋ] is a non-contact variant

- It seems clear that this is due to language contact in the sense of Thomason (2008)

- Further research is needed to establish this without a shadow of a doubt
(cf. Poplack & Levey 2010)

Survey

- Conducted in Ho Chi Minh City, 2015
    - Word list (n=50)
    - Sentences (n=50)
    - Consultants asked to speak as they do at home
        - → Colloquial register

Results: Gradience

- Intra-speaker variation
- Speaker MXL:
  - [jːàʔ] 'market' (Word List)
  - [rːàʔ] 'market' (Sentence)

	Non-contact	Both contact/ non-contact	Only contact
/r/,/j/	2	21	7

	Non-contact	Both contact/ non-contact	Only contact
/ŋ/ / V$_{rd}$ _	∅	28	2

Results: Gradience

- Variation is not obviously lexical diffusion

Word	Non-contact	Contact	Total
/rːàʔ/ 'market'	20	16	36
/kra/ 'monkey'	28	25	53
/krɔŋ/ 'river'	28	20	48
/hrɛj/ 'today'	16	21	37
/kyɔw/ 'wood'	19	25	44
/pːròj/ 'yesterday'	26	29	55

Results: Gradience

- Variation is not obviously lexical diffusion

Word	Non-contact	Contact	Total
/vɔ̀ŋ/ 'circle'	24	3	27
/ʔɪŋ ʔɔ̀ŋ/ 'frog'	36	18	54
/thɔ̀ŋ/ 'knife'	6	51	57
/krɔŋ/ 'river'	24	35	59
/pùŋ/ 'top.of'	28	28	56
/thɔŋ/ 'with'	5	38	43

Results: Gradience

- Not obviously a lexical implicational hierarchy
- If a speaker only uses a variable in one lexical item:

[r]	[j]/[z]	[ŋ]	[ŋ͡m]
/krɔŋ/ 'river'	/kyɔw/ 'wood'	/vɔ̀ŋ/ 'circle'	/thɔ̀ŋ/ 'knife'
/pːròj/ **'yesterday'**	/kyɔw/ 'wood'	/ʔɪŋ ʔɔ̀ŋ/ 'frog'	/thɔŋ/ **'with'**
	/pːròj/ **'yesterday'**	/pùŋ/ 'top.of'	
		/thɔŋ/ **'with'**	

Results: Village

- Likelihood ratio tests for /r/ and /ŋ/ (cf. Baayen 2008)
    - Random effects: Speaker, Lexical item
    - Factors: Age, Gender, Task (e.g. Word List), Village

- /r/: Village significant ($p < 0.05$)
    - Age, Gender, Task n.s.
- /ŋ/: Village significant ($p < 0.05$)
    - Age, Gender, Task n.s.

Results: Village

- Two (of 8) villages predicted [j]/[z] & [ŋm͡]: Palei Hamu Tanran & Palei Hamu Craok

- Baclawski (2016) claims that these two villages have greater socioeconomic contact with Vietnamese communities

Interim summary

- Two phonological contact variables

- Gradient inter- and intra-speaker variation

- Significance of speaker village, but much variation left to be explained

Syntactic variable

- Optional *wh*-topicalization
  - Licit for some Eastern Cham speakers
  - Ungrammatical in Vietnamese

(1)   {%}   *zut*   *ʔːa*   {*thɛj*} *băŋ*   *lːɔ*   *mːɔ*
                 friend   invite   who   eat   meat   cow
    'Who did you [friend] invite to eat beef?'   EASTERN CHAM

(1')   {*}   *Bạn*   *mời*   {*ai*}   *ăn*   *thịt*   *bò?*
                 friend   invite   who   eat   meat   cow
    'Who did you [friend] invite to eat beef?'   VIETNAMESE

Information structure & syntax

- Topic- (and focus-)movement is seen as the result of prosodic factors, or features in the lexicon

- Optional 'p-movement'/prosodic movement
  - Zubizarreta (1998), Szendrői (2001), Horváth (2010), Samek-Lodovici (2015), Richards (2016), etc.

- [topic] features in the syntax
  - Rizzi (1997)'s Cartographic approach, López (2009), etc.
  - Optionally assigned (Mikkelsen 2005; cf. Bolinger 1972)

Information structure & syntax

- Topic feature in the syntax
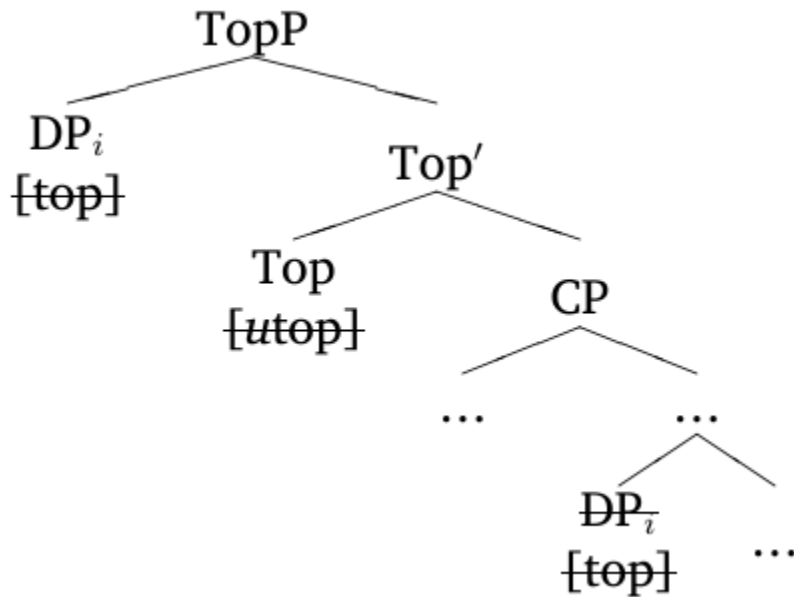
- Prosodic constraint/rule: e.g. topics must be in a separate intonational phrase
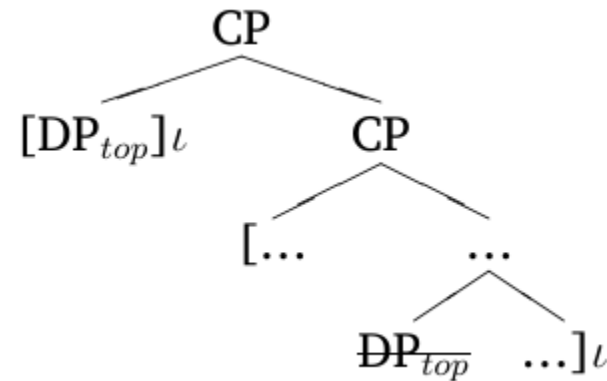
Topic as discourse anaphora

- There are many notions of 'topic'
  - A. Aboutness
  - B. Old information (cf. Reinhart 1981)
  - C. Discourse relevant

- López (2009) discards A,B
  - Supports a 'discourse subordination test' (Asher & Lascarides 2002; Asher & Vieu 2005)

Topicalization

- Discourse coordination context ('X, then Y'):
  - Catalan: ✖ clitic right-dislocation
  - Eastern Cham: ✖ topicalization

CONTEXT: 'He cooked the meat.'
(2) #*Després   se   la   va   menjar,   **la   carn.**
    afterwards   CL   CL.ACC   PAST   eat.PART   the   meat
    'Afterwards, he ate the meat.' (López 2009: 48)   CATALÁN

CONTEXT: 'My older brother cooked chicken.'
(2') #*plɔh năn,   l:ɔ   n:ŭʔ,   aj   băŋ*
     after.that   meat   chicken   o.s.   eat
     'After that, my older brother ate the chicken.'   EASTERN CHAM

Topicalization

- Discourse subordination ('X, because Y'):
    - Catalan: ✓clitic right-dislocation
    - Eastern Cham: ✓ topicalization

CONTEXT: 'Look at him cooking the meat.'
(3) *La*   *fa*   *molt*   *be,*   *el Joan,*   **la carn**.
    CL.ACC make   very   well   the Joan   the meat
    'He [Joan] cooks the meat very well.' (López 2009: 49)   CATALÁN

CONTEXT: 'Look at him cutting the meat.'
(3') **rǎlɔ**,   *ɲu sĭʔ*   *ʃjɛm*   *lo*
    meat   3SG chop   good   very
    'He cuts the meat very well.'   EASTERN CHAM

Wh-topicalization

- The discourse subordination test also applies to (D-linked) *wh*-phrases in Eastern Cham
  - cf. Pan (2016) on Mandarin

CONTEXT: A: 'I ate dinner.'
(4) #*hwăʔ* *plɔh*, *ƙeit*, *hɨ* *ŋăʔ*
    eat.rice after what 2SG do
    B: 'After eating, what did you do?' [Coordination with A]

(4') *ƙeit*, *zut* *băŋ*
    what friend eat
    B: 'What [kind of food] did you eat?' [Subordination with A]

Topicalization in Vietnamese

- Topicalization is marked overtly with *thì*
- *Wh*-topicalization seems to be impossible

CONTEXT: A: 'I ate dinner.'

(4) #*Sau khi   ăn tối,   (cái)   gì   (thì)   bạn   làm?*
     after   eat.dinner   CLF   what   TOP   friend   do
     B: 'After eating dinner, what did you do?'   [Coordination with A]

(4') #*(Cái)   gì   (thì)   bạn   ăn?*
      CLF   what   TOP   friend   eat
      B: 'What [kind of food] did you eat?'   [Subordination with A]

Syntax/discourse elicitation

- Speaker information (7 speakers)
- Elicitation of 4 discourse coordination,
  4 discourse subordination contexts

Context	Discourse coordination	Discourse subordination
A: 'First, I cooked dinner.'	B: 'Then, **what** did you do?'	B: '**What** did you cook?'
A: 'First, I invited someone.'	B: 'Then, **who** did Kenny invite?'	B: '**Who** did you invite?'
A: 'I ate dinner.'	B: '**What** did you eat?' [Clarification]	B: '**What** did you eat?' [Elaboration]
A: 'She isn't eating meat.'	B: '**What** does she like to eat?'	B: 'Then, **what** is she eating?'

Variation in grammaticality

- 5 speakers accepted *wh*-topicalization
- 2 speakers categorically rejected it

CONTEXT: A: 'First, I invited someone [to come eat beef with us].'
(5) *plɔh,*   ***thɛj,***   *zut*   *ʔ̥:a*   [Subordination with A]
     before   who   friend   invite
     B: 'Who did you [friend] invite, then?' (Speaker NNA)

(5') #***thɛj,***   *zut*   *ʔ̥:a*   *bǎŋ l:ɔ*   *m:ɔ*   [Subordination with A]
      who   friend   invite   eat meat   cow
      B: 'Who did you [friend] invite to eat beef?' (Speaker NTNT)

Overgeneralization

- 2 speakers accepted *wh*-topicalization in all contexts (but did not necessarily produce it)

CONTEXT: A: 'I just ate dinner.'
(5) k̥ɛɪt,   jut   p̥ja hwă?   [Coordination with A]
    what   friend   just eat.rice
    B: 'What did you [friend] eat? [Clarification]' (Speaker TQD)

CONTEXT: A: 'That woman is cooking duck.'
(6) k̥i   p̥ɔ nĭ? m:ɛj   năn ̥tɔ?   ŋă?   [Coordination with A]
    what   REL CLF woman   that PROG   make[cook]
    B: 'What is that woman cooking? [Clarification]' (Speaker DPNS)

Results: Syntactic variation

- Varying acceptability of *wh*-topicalization

	Group 1a			Group 1b		Group 2	
	NNA	ER	TDK	TQD	DPNS	TTL	NTNT
wh-topicalization	✓	✓	✓	✓*	✓*	✘	✘
Contact with VN?	✘	✘	✘	?	?	✓	✓

Covariance

- There is a weak covariance between the phonological and syntactic variables

	Group 1a			Group 1b			Group 2	
	NNA	ER	TDK	TQD	DPNS	Average	TTL	NTNT
/r/,/j/ → [j]/[z]	18%	36%	100%	50%	80%	57%	66%	100%
/ŋ/ → [ŋ͡m]	35%	47%	5%	30%	60%	59%	100%	74%
wh-topicalization	✓	✓	✓	✓*	✓*		✘	✘

Covariance

- There is a weak covariance between the phonological and syntactic variables:
  - If *wh*-topicalization is ungrammatical, speakers predominantly use contact phonological variants
  - **If non-contact phonological variants are predominant, *wh*-topicalization is grammatical**

	Group 1a			Group 1b			Group 2	
	NNA	ER	TDK	TQD	DPNS	Average	TTL	NTNT
/r/ → [j]/[z]	18%	36%	100%	50%	80%	57%	66%	100%
/ŋ/ → [ŋ͡m]	35%	47%	5%	30%	60%	59%	100%	74%
wh-topicalization	✓	✓	✓	✓*	✓*		✘	✘

Covariance

- There is a weak covariance between the phonological and syntactic variables:
  - Group 2: More contact in syntax → more contact in phonology
  - **Group 1a: Less contact in phonology → less contact in syntax**

	Group 1a			Group 1b			Group 2	
	NNA	ER	TDK	TQD	DPNS	Average	TTL	NTNT
/r/ → [j]/[z]	18%	36%	100%	50%	80%	57%	66%	100%
/ŋ/ → [ŋ͡m]	35%	47%	5%	30%	60%	59%	100%	74%
wh-topicalization	✓	✓	✓	✓*	✓*		✘	✘

Covariance & salience

- Salience could be a crucial factor
- Diglossia is very dominant in language attitudes, could mask language contact effects
- These non-diglossic phonological variables are conspicuously non-salient
  - No speakers have offered metalinguistic commentary, even when confronted with the forms (Baclawski 2016)
- The syntactic variable is likewise non-salient

An asymmetry

- *Wh*-topicalization is always optional, possibly obscuring speaker salience
- Group 2 speakers never utter infelicitous utterances to Group 1
- *Wh*-topicalization is infelicitous to Group 2

Discourse subordination context	Group 1	Group 2
wh-in-situ	✓	✓
wh-topicalization	✓	✘

Covariance & salience

- Perhaps all these contact effects are due to attitudes toward bilingualism
  (cf. Matras 2009)
  - Group 1a inhibits replication of Vietnamese syntax and, for a subset of speakers, phonology
  - Group 2 does not inhibit replication of Vietnamese syntax or phonology

The continuing search for covariance

- At least for intense language contact situations, syntactic variation can covary with phonological variation

- Much future research is needed:
  - Perception studies to assess salience
  - Larger sample size for discourse elicitation
  - Greater study of Eastern Cham and Vietnamese variation and contact

Acknowledgments

- Thank you very much!

- I would like to acknowledge all the Cham people, my research assistant Hamu Ligaih (Đàng Thanh Quốc Thuận) and his professor, Sakaya (Dr. Trương Văn Món) for their integral role in translation, finding speakers, and the general organization of the study

- Additional thanks to my Cham teacher, Mohammad Saleh Thiên. Thanks to Andrew Garrett, Keith Johnson, Line Mikkelsen, Peter Jenks, Terry Kaufman, Jim Matisoff, Justin Davidson, Johanna Nichols, Marc Brunelle, Spencer Lamoureux, Amalia Skilton, Elise Stickles, and Anna Jurgensen; special thanks to my research apprentices, Win Htet Kyaw and Nathan Phillip Cahn; and to my trusty fieldwork assistant, Dylan Calhoun