

# Temporal Links and L2 Development: An Application of NLP Analysis

- Jeff Connor-Linton & Inderjeet Mani  
Department of Linguistics  
Georgetown University

[connorlj@georgetown.edu](mailto:connorlj@georgetown.edu)

[im5@georgetown.edu](mailto:im5@georgetown.edu)

# Abstract

- Temporal Links and L2 Development: An Application of NLP Analysis
- Most research on L2 development has focused on the learner's approximation of native-like use of a few linguistic forms. Some recent SLA research analyzes learner's interlanguages across different modalities of production, using relatively gross measures of syntactic or lexical complexity. Even multivariate analyses of cooccurrence relations among linguistic features are constrained by their reliance on frequency measures of directly observable lexico-grammatical features. So far, few analytic tools have been developed to represent levels of discourse structure (e.g., cohesion), which rely more on links between elements than frequencies of various forms. Recent progress by computational linguists enables us to represent aspects of narrative structure in text, by analyzing temporal links between events and time expressions mentioned in a text. This allows us in turn to compare patterns of temporal links across stages of L2 development.
- This paper presents an exploratory analysis of temporal links in English oral and written proficiency samples elicited from 16 speakers of English: 8 native speakers and 8 students in 'Advanced' courses in an Intensive English Program. This pilot corpus includes 5888 words elicited from subjects via a written narrative retelling task.
- The paper describes similarities and differences across learner groups in the sample--not only the number of events, temporal expressions and their textual links, but the characteristics of the webs of temporal relations constructed by learners in comparison with those constructed by native speakers.
- More generally, the paper will demonstrate the potential value of applying analytic procedures developed for natural language processing purposes to the study of second language acquisition.

# TimeML Annotation

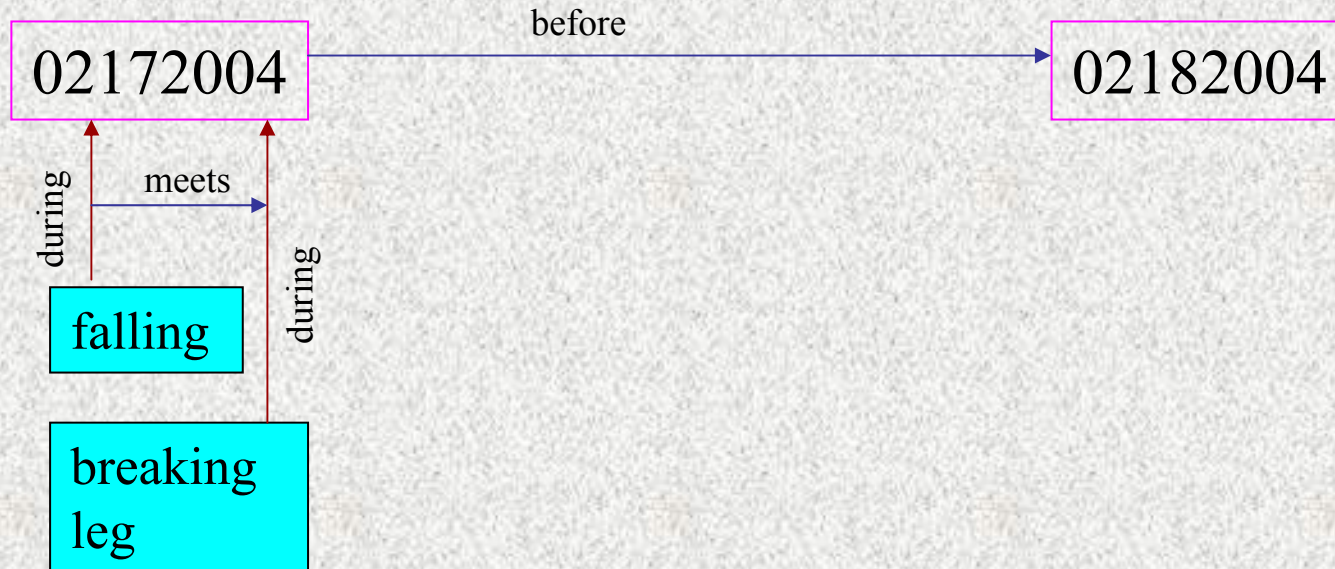
- TimeML is a proposed metadata standard for markup of events and their temporal anchoring and ordering
- Consists of EVENT tags, TIMEX3 tags, and LINK tags
  - EVENTS are grouped into classes and have tense and aspect features
  - LINKS include overt and covert links
    - Can be within or across sentences

# TLINK

Expresses temporal ordering between events and/or times, using fourteen possible temporal relations

- *before, immediately before, simultaneous, identical, includes, begins, ends, and inverses*

Yesterday Holly fell and broke her leg.



# SLINK and ALINK

- SLINK is used for contexts introducing subordination relations
  - **Modal:** Relation introduced mostly by modal verbs (**should, could, would, etc.**) and events that introduce a reference to a possible world --mainly I\_STATES:
    - *Mary wanted John to buy some wine.*
  - **Factive:** Certain verbs introduce an entailment (or presupposition) of the argument's veracity. They include **forget** in the tensed complement, **regret, manage:**
    - *John forgot that he was in Boston last year.*
  - *etc.*
- ALINK or Aspectual Link represent the relationship between an aspectual event and its argument event.
  - **1. Initiation:**
    - *John started to read.*
  - **2. Culmination:**
    - *John finished assembling the table.*
  - **3. Termination:**
    - *John stopped talking.*
  - **4. Continuation:**
    - *John kept talking.*

# Annotation Tools

- Callisto      [callisto.mitre.org](http://callisto.mitre.org)
  - java-based annotation tool to support linguistic annotation of textual sources for any Unicode-supported language
- TANGO      [www.timeml.org](http://www.timeml.org)
  - java-based annotation tool that provides a graphical interface for annotation

# Callisto Shot

Callisto - J:\Documents and Settings\IMANI\Desktop\AAAL04\E2s/e632awn.txt.tml.u.tml

File Edit Format Tools Help

E63A1SW-A.WN

After getting out of jail, which seemed not to please at all, Charlie Chaplin looked for a job. Eventually, he was hired in a company which i was building boats. Full of enthusiasm, he was ordered to find a piece of wood , but he took the wrong one, the piece of wood that was preventing the boat from sinking. Therefore, he was naturally fire. Meanwhile, the girl who stole bannanas in the port (let's call her Anna) discovered her father laying on the street, killed by the police during a manifestation. The police came to her house to take care of her and her two little sisters but, taking advantage of a moment of inattention of the policemen, she escaped, not willing to spend her life in a special establishment. However, time was hard and she had to struggle to find something to eat.

While she was passing by a bakery, she stole a bread but, unfortunately an old woman saw her stealing it and called out the baker. As she was running away, she collided with Chaplin. The baker grabbed her hand and called the police but Chaplin stopped him by saying that he was the guilty person.

While the policeman look him away, the old woman went back at the baker and asserted that the girl was guilty. Chaplin, who definitely wanted to return in jail, Decided to have lunch in a cafeteria knowing that he wouldn't be able to pay. He eventually ended up in the police car with other people, especially Anna who was also caught. On their way to jail, the police car crashed while Anna was trying to escape. They collapsed on the street, her, Chaplin and another policeman. She ran away, convincing a doubtful Charlie Chaplin to join her. They observed as they were chatting, the life of a successful family who could make it on the other end. Finally, Anna found out a "house", which was more of a 10 m2 100 years old refuge. They, finally, had a home and was even able to imitate the successful family's way of life.

EVENT Tag

Timex3 Tag

Signal Tag

Delete Annotation

Events Signals Timex3 ALinks SLinks TLinks

Event ID	Class	Tense	Aspect
(e1) getting	OCCURRENCE	PAST	PROGRESSIVE
(e43) to please	I_STATE	NONE	NONE
(e2) looked	OCCURRENCE	PAST	NONE
(e3) was hired	OCCURRENCE	PAST	NONE
(e4) was building	OCCURRENCE	PAST	PROGRESSIVE
(e5) was ordered	OCCURRENCE	PAST	NONE
(e44) to find	OCCURRENCE	NONE	NONE
(e6) took	OCCURRENCE	PAST	NONE
(e7) waspreventing	OCCURRENCE	PAST	PROGRESSIVE
(e45) sinking	OCCURRENCE	PAST	PROGRESSIVE
(e8) was naturally fire	OCCURRENCE	PAST	NONE

Font: 12pt. dialog Charset: UTF-8 Task: Tango TimeML

# TANGO Shot

Tango - C:\Documents and Settings\IMANI\Desktop\AAAL04\2s\le632awn.txt.tml.u.tml

File

E63A1SW-A.WM

After *getting* out of jail, which seemed not *to please* at all, Charlie Chaplin *looked* for a job. Eventually, he *was hired* in a company which i *was building* boats. Full of enthusiasm, he *was ordered to find* a piece of wood , but he *took* the wrong one, the piece of wood that *was preventing* the boat from *sinking*. Therefore, he *was naturally fire*. Meanwhile, the girl who *stole* bannanas in the port (let's call her Anna) *discovered* her father *laying* on the street, *killed* by the police during a manifestation. The police *came* to her house *to take* care of her and her two little sisters but, *taking advantage* of a moment of inattention of the policemen, she *escaped*, not willing *to spend* her life in a special establishment. However, time was hard and she *had to struggle* to find something to eat.

View XML View Prefs Show Grid Hide SLinks Hide TLinks Hide ALinks User Links Hide Unkn... Score GrappaView

Select Delete Events S-Links T-Links A-Links SmartLink Closure Arrange

A-Z  
1-3

# Closure Tool

- Automatically expands TLINKs using axiomatic temporal reasoning
- Uses transitivity relations among TLINKS
  - e.g., e1 BEFORE e2 & e2 INCLUDES e3  
=> e1 before e3
- Integrated with TANGO

# Event Frequencies - 2

Events: Tense					
E1				E2	
	%				%
152	84%	PRESENT		22	14%
14	8%	PAST		125	77%
13	7%	NONE		15	9%
1	1%	FUTURE			

# TLinks - 1

Tlinks (Unclosed)					
E1			E2		
	%			%	
298	79%	BEFORE	169	60%	
19	5%	INCLUDES	8	3%	
18	5%	IBEFORE	6	2%	
11	3%	SIMULTANEOUS	11	4%	
6	2%	IDENTITY	14	5%	
6	2%	OVERLAP_BEFORE	1	0%	
5	1%	OVERLAP_AFTER			

# TLinks - 2

E1	%		E2	%
5	1%	IS_INCLUDED	5	2%
5	1%	BEGINS	3	1%
3	1%	ENDS	4	1%
1	0%	AFTER	44	16%
1	0%	DURING	10	4%
1	0%	ENDED_BY	1	0%
		BEGUN_BY	3	1%
		IAFTER	1	0%

# Complexity Metrics - 1

Complexity Metrics				
	Wds/S	Wds/Event	Wds/TLink	TLinks/Event
E1 mean	9.5	5.9	8.2	0.73
E1 SD	1.6	0.5	1.1	0.07
E2A Mean	8.7	6.6	10.1	0.66
E2A SD	1.2	1.1	1.8	0.06
E1:E2A t	0.30	0.12	0.03	0.05

# Closure Tool

- Automatically expands TLINKs using axiomatic temporal reasoning
- Uses transitivity relations among TLINKS
  - e.g., e1 BEFORE e2 & e2 INCLUDES e3  
=> e1 before e3
- Integrated with TANGO

# Pre-Closure

Tango - C:\Documents and Settings\IMANI\Desktop\AAAL04\E2s\E662AWN.txt.i.tml

**File**

The second half of the movie *starts* with the scene of the street where the father of the beautiful girl and their little siblings *lies* dead. The charming girl *runs* to her father and *tells* the policeman that that is her father.

The place *is changed* to the girl's house for the next scene. One policeman and the girl *are consoling* two little children, and the other two policemen -- or maybe detectives -- *are sitting* at the table and *discussing* her father's death. In the meanwhile, she *sneaks* out of the house, and the policeman were surprised *to see* where she *'s gone*.

Charlie Chaplin who *got* credit of *preventing* some other prisoners from *exiling* the prison in the first half of the movie, *is now talking* with a gentleman, who *looks* like the top administrator of

View XML View Prefs Show Grid Hide SLinks Hide TLinks Hide ALinks User Links... Hide Unkn... Score GrappaView

Select Delete Events S-Links T-Links A-Links SmartLink Closure Arrange

A-Z  
1-3

# Post-Closure

Tango - C:\Documents and Settings\IMAN\Desktop\AAAL04\E2s\E662AWN.txt.i.closed.tml

**File**

*consoling* two little children, and the other two policemen -- or maybe detectives -- *are sitting* at the table and *discussing* her father's death. In the meanwhile, she *sneaks* out of the house, and the policeman were surprised *to see* where she's *gone*.

Charlie Chaplin who *got* credit of *preventing* some other prisoners from *exiling* the prison in the first half of the movie, *is now talking* with a gentleman, who *looks* like the top administrator of the prison. And then, he *gives* a letter over to Charlie.

With the letter, Charlie is at the boat construction site. He *hands* the letter over the a guy who seems like a responsible man at the place, and *takes* off the jacket and the hat and *puts* down the stick, and then *starts to work*.

View XML View Prefs Show Grid Hide SLinks Hide TLinks Hide ALinks User Links... Hide Unkn... Score GrappaView

Select Delete Events S-Links T-Links A-Links SmartLink Closure Arrange

A-Z  
1-3

# Closed Links: A Measure of Connectivity

Closed TLinks: A Measure of Connectivity		
	E1	E2A
Tlinks	12330	4924

# Conclusion

- Differences between native speakers (NS) and non-native speakers (NNS) approach significance (at  $p < .05$  level) when comparing NS/NNS on TLinks/Events and actually are significant ( $p = .03$ ) on Wds/TLink.
  - On average, NSs use fewer wds to create TLinks (8.2/TLink vs. 10.1 for NNSs).
- A slightly higher proportion of NS Events are TLinked (.73 vs .66 for NNSs;  $p = .053$ )
- Number of closed TLINKS for NS far exceeds the number for NNS (12,330 vs. 4924).
  - This means NS have, on the average, longer chains of TLINKS

# BACKUPS

# Link Frequencies

All Links (Unclosed)				
E1	%		E2	%
516		EVENT	427	
496		LINKS	378	
379	76%	TLINK	280	74%
11	2%	ALINK	13	3%
106	21%	SLINK	85	22%

# Event Frequencies - 3

Events: Aspect				
E1	%		E2	%
143	79%	NONE	124	77%
30	17%	PROGRESSIVE	30	19%
6	3%	PERFECTIVE	7	4%
1	1%	PERF-PROG	1	1%

# Event Frequencies - 1

Events: Class				
E1	%		E2	%
138	77%	OCCURRENCE	117	72%
11	6%	I_ACTION	3	2%
10	6%	PERCEPTION	11	7%
7	4%	STATE	7	4%
7	4%	REPORTING	4	2%
5	3%	ASPECTUAL	11	7%
2	1%	I_STATE	5	3%
		MODAL	4	2%

# SLinks

<b>Slinks (Unclosed)</b>					
E1	%		E2	%	
81	76%	MODAL	62	73%	
21	20%	EVIDENTIAL	22	26%	
3	3%	FACTIVE			
1	1%	COUNTER_FACTIVE			
		NEG_EVIDENTIAL	1	1%	

# ALinks

<b>ALinks (Unclosed)</b>				
<b>E1</b>	<b>%</b>		<b>E2</b>	<b>%</b>
6	55%	INITIATES	8	62%
3	27%	CONTINUES	3	23%
2	18%	CULMINATES		
		TERMINATES	2	15%

# Units of Measurement - 1

Units of Measurement				
	Words	Sentences	Events	TLinks
E1	3085	343	516	379
E2A	2803	323	427	280

# Units of Measurement - 2

	<b>Units: Mean Frequencies</b>			
	Words	Sentences	Events	TLinks
E1 mean	385.6	42.9	65.3	47.3
E1 SD	143.0	18.7	23.1	16.9
E2A Mean	350.4	40.4	54.8	35.0
E2A SD	165.7	16.8	27.3	14.9
E1:E2A t	0.67	0.79	0.44	0.16

# Units of Measurement - 3

	r
#Words:#Sentences	0.94
#Words:#Events	0.97
#Words:#Tlinks	0.90
#Sentences:#Events	0.94
#Sentences:#Tlinks	0.9

# Complexity Metrics - 2

	r
Wds/S:Wds/Event	0.22
Wds/S:Wds/Tlink	0.35
Wds/S:Tlinks/Event	-0.32
Wds/Event:Tlinks/Event	-0.2
Wds/Tlink:Tlinks/Event	-0.67