

MT Extravaganza

Ling 567

March 10, 2010

Overview

- Background
- Overview results
- Interactive exploration
- Course evals

Languages

iso	Language name	Language family	Grammar authors
bre	Breton	Indo-European	Foster & Todd
chr	Cherokee	Iroquoian	Hold & Wu
eng	English	Indo-European	
fra	French	Indo-European	Rolston
ita	Italian	Indo-European	
jaa	Jamamadí	Arauan	Cross & Wax
lut	Lushootseed	Salishan	Adam-Collier & Mills
ojg	Nishnaabemwin	Algic	Klasswn & Shaw
pli	Pali	Indo-European	Frowein
pbt	Pashto	Indo-European	Kauhl & Scott
rus	Russian	Indo-European	Gracheva & Suskic
sna	Shona	Niger-Congo	Crowgey & Pham
vie	Vietnamese	Austro-Asiatic	Song & Shu

Grammar coverage

- Basic word order
- Case
- Agreement
- Personal pronouns
- Tense/aspect
- Sentential negation
- Argument optionality
- Matrix yes-no questions
- Coordination
- Modification (adjective, adverb)
- Non-verbal predicates
- Clausal complements
- Information structure

Set up

- Transfer-based MT: Grammars parse and generate, mapping surface strings to semantic representations in MRS
- Quasi lexical interlingua (English lemmatas as PRED values)
- ‘semi’ (Semantic Interface) maps variable properties (PNG, TAM, COG-ST, INFO-STR) from grammar internal space to interlinual space. Lossy mapping, provides defaults.
- One ‘accommodation’ transfer grammar per language, instantiating shared transfer rules.

Input sentences

- | | | | |
|-----|---|-----|-------------------------------------|
| 1. | Dogs sleep | 12. | Dogs chase cars and cats chase dogs |
| 2. | Dogs chase cars | 13. | Cats chase dogs and sleep |
| 3. | I chase you | 14. | Do cats chase dogs? |
| 4. | These dogs sleep | 15. | Hungry dogs eat |
| 5. | Dogs eat | 16. | Dogs eat quickly |
| 6. | I can eat glass | 17. | The dogs are hungry |
| 7. | It doesn't hurt me | 18. | The dogs are in the park |
| 8. | The dogs chase cars | 19. | The dogs are the cats |
| 9. | I think that you know that dogs chase cars | 20. | The dogs-FP chase the cats |
| 10. | I ask whether you know that dogs chase cars | 21. | The dogs chase-FP the cats |
| 11. | Cats and dogs chase cars | 22. | The dogs chase the cats-FP |

Items with end to end success													
	bre	chr	eng	fra	ita	jaa	lut	ojg	pbt	pli	rus	sna	vie
bre	—	4	16	13	15	3	10	14	15	11	12	12	0
chr	6	—	6	5	7	4	5	7	7	6	7	7	0
eng	15	8	—	17	18	4	11	16	18	13	16	15	16
fra	12	4	18	—	16	3	4	14	15	11	16	10	0
ita	9	7	13	9	—	3	7	13	9	9	9	11	0
jaa	3	3	15	11	15	—	3	12	12	10	1	12	0
lut	13	5	13	12	13	3	—	8	9	10	10	4	0
ojg	7	3	17	9	16	4	3	—	13	12	15	9	0
pbt	14	5	18	14	18	4	4	14	—	13	15	13	0
pli	12	3	14	12	13	2	8	11	10	—	11	9	0
rus	7	3	9	8	8	2	4	7	9	7	—	8	0
sna	6	3	9	11	9	2	3	7	8	10	11	—	0
vie	13	4	21	15	15	1	13	10	14	10	14	4	—

Items with exact match output													
	bre	chr	eng	fra	ita	jaa	lut	ojg	pbt	pli	rus	sna	vie
bre	—	3	11	6	8	3	3	3	9	4	4	4	0
chr	3	—	4	4	3	3	1	0	4	4	3	2	0
eng	6	2	—	11	12	4	6	2	13	9	6	8	0
fra	3	2	9	—	7	3	3	1	8	8	6	4	0
ita	4	2	6	6	—	3	3	1	6	5	1	2	0
jaa	2	1	6	6	7	—	1	1	6	4	1	2	0
lut	2	1	6	1	4	0	—	0	2	3	2	1	0
ojg	3	2	7	5	9	4	1	—	4	4	2	5	0
pbt	4	2	13	9	8	4	2	0	—	6	7	6	0
pli	4	1	7	9	8	2	3	1	6	—	3	3	0
rus	2	2	8	7	3	1	1	2	8	5	—	6	0
sna	2	2	6	6	4	1	1	2	3	3	4	—	0
vie	2	1	8	4	2	0	4	1	6	2	4	2	—

Average number of outputs, if any													
	bre	chr	eng	fra	ita	jaa	lut	ojg	pbt	pli	rus	sna	vie
bre	–	11.25	5.13	3.31	7.93	2.0	3.9	15.64	6.2	7.09	32.42	4.25	0
chr	9.0	–	4.0	4.0	9.29	2.0	6.4	3.71	1.71	7.67	6.14	4.29	0
eng	18.87	77.88	–	1.47	4.06	2.5	2.45	11.81	6.61	5.15	26.19	2.33	0
fra	25.42	66.25	9.22	–	15.5	2.0	3.5	7.5	2.2	5.18	13.63	1.2	0
ita	26.33	90.71	63.77	1.44	–	2.0	2.29	10.08	1.44	4.11	94.0	2.82	0
jaa	5.0	17.67	57.0	3.73	18.07	–	2.0	4.58	3.08	6.0	4.0	4.67	0
lut	27.23	65.0	33.85	42.0	8.54	9.67	–	36.75	40.44	16.5	28.9	12.75	0
ojg	27.57	9.67	37.76	3.78	36.19	2.0	1.67	–	9.38	7.25	44.93	2.78	0
pbt	26.36	34.2	7.89	1.36	20.22	2.25	9.0	17.21	–	5.62	17.4	3.31	0
pli	9.83	38.0	8.57	1.25	1.15	2.0	2.25	5.73	1.4	–	25.45	1.11	0
rus	15.43	39.0	6.0	1.88	18.75	4.0	41.25	4.14	2.22	6.57	–	1.63	0
sna	23.33	39.0	9.78	2.55	10.44	6.0	1.67	3.14	5.13	6.2	16.09	–	0
vie	34.46	19.75	13.95	9.27	22.0	6.0	3.31	5.2	10.07	30.8	32.36	2.0	–

Maximum number of outputs													
	bre	chr	eng	fra	ita	jaa	lut	ojg	pbt	pli	rus	sna	vie
bre	—	20	16	16	80	2	12	96	32	16	70	10	0
chr	24	—	8	8	30	2	14	8	3	24	12	18	0
eng	144	216	—	3	30	4	6	56	56	16	162	14	0
fra	96	152	48	—	81	2	8	48	7	10	54	2	0
ita	144	252	416	3	—	2	6	48	4	6	288	20	0
jaa	6	28	256	16	54	—	2	32	8	12	4	36	0
lut	100	252	144	225	20	12	—	56	128	30	64	36	0
ojg	112	16	384	11	252	2	2	—	51	16	144	8	0
pbt	144	89	20	3	144	3	32	56	—	18	54	18	0
pli	63	86	48	2	3	2	6	21	4	—	96	2	0
rus	32	89	9	3	80	6	160	8	4	16	—	2	0
sna	48	89	24	4	30	6	2	6	8	16	36	—	0
vie	144	28	112	24	96	6	10	9	48	60	96	2	—

Successful source languages per item, items 1 to 11											
tgt	1	2	3	4	5	6	7	8	9	10	11
bre	12	9	12	8	5	0	9	8	0	0	6
chr	12	3	11	8	6	9	0	1	0	0	0
eng	12	11	10	8	11	9	9	10	4	2	9
fra	12	11	5	8	9	6	6	10	1	0	8
ita	12	11	11	8	11	9	9	10	5	1	9
jaa	11	0	11	1	7	1	0	0	0	0	0
lut	12	5	11	8	3	0	8	6	0	0	0
ojg	12	11	11	1	11	3	7	10	0	0	5
pbt	12	11	7	8	3	9	6	10	1	1	9
pli	12	12	12	7	11	0	9	11	0	0	0
rus	12	11	10	1	11	8	7	10	0	0	5
sna	12	9	10	9	8	8	8	9	3	1	2
vie	1	1	1	1	0	0	1	1	1	1	0

Successful source languages per item, items 12 to 22											
tgt	12	13	14	15	16	17	18	19	20	21	22
bre	3	5	7	2	0	5	2	4	9	4	7
chr	0	0	2	0	0	0	0	0	0	0	0
eng	5	9	9	8	6	4	1	5	10	9	8
fra	7	6	8	6	0	0	0	6	10	9	8
ita	6	8	9	8	6	0	0	0	11	10	9
jaa	0	0	0	0	0	0	1	3	0	0	0
lut	0	0	5	1	1	5	0	0	4	3	3
ojg	3	7	9	8	5	0	0	4	10	8	8
pbt	5	9	9	1	1	5	0	5	10	9	8
pli	0	0	9	8	4	0	0	0	10	9	8
rus	1	7	9	9	6	4	1	1	9	8	7
sna	0	0	7	0	0	0	0	6	8	7	7
vie	1	1	1	0	0	1	1	1	1	1	0

Successful target languages per item , items 1 to 11											
src	1	2	3	4	5	6	7	8	9	10	11
bre	11	9	10	8	9	0	9	9	4	0	6
chr	11	9	9	11	11	6	0	10	0	0	0
eng	12	11	12	9	8	6	10	11	4	3	7
fra	11	8	11	8	8	7	0	8	2	1	6
ita	11	10	9	8	8	6	3	9	1	1	6
jaa	11	7	8	0	6	5	6	7	1	0	4
lut	11	9	10	8	9	0	6	0	0	0	4
ojg	11	5	9	0	7	7	9	5	0	0	4
pbt	11	8	11	8	8	7	8	8	1	0	6
pli	11	9	1	8	8	0	1	9	1	0	5
rus	11	8	11	0	0	6	9	8	0	0	0
sna	11	3	11	0	6	6	8	3	0	0	0
vie	11	9	10	8	8	6	10	9	1	1	5

Successful target languages per item, items 12 to 22											
src	12	13	14	15	16	17	18	19	20	21	22
bre	5	6	9	6	0	0	0	6	9	0	9
chr	0	0	0	0	0	0	0	0	0	0	0
eng	6	6	11	5	4	5	2	7	10	9	9
fra	0	5	8	5	5	5	1	0	8	8	8
ita	4	4	10	5	4	0	0	0	0	0	0
jaa	4	4	7	0	0	1	0	5	7	7	7
lut	3	5	0	7	0	4	0	0	8	8	8
ojg	3	4	5	5	4	0	1	5	8	8	8
pbt	3	6	8	6	4	0	0	6	8	7	8
pli	1	7	9	5	3	0	0	1	9	9	8
rus	0	0	0	0	0	4	0	0	8	7	0
sna	0	0	8	0	0	0	0	0	8	7	8
vie	2	5	9	7	5	5	2	5	9	7	0

Overview

- Background
- Overview results
- Interactive exploration
- Course evals