Developing a Japanese Corpus Annotated for "Frames" and their "Elements" –Specifying What People Understand with MSFA–

Kow KURODA National Institute of Communications Technology (NICT), Japan 09/06/2006

Today's Topic

- Introducing *Multi-layered/dimensional Semantic Frame Analysis* (MSFA: Kuroda and Isahara 2005)
 - It was developed as an annotation scheme hopefully compatible with *Berkeley FrameNet* (henceforth, BFN) (Baker, et al. 1998; Fillmore, et al. 2003; Johnson and Fillmore 2001; Lowe, et al. 1997)

• Caveats:

- So far, MSFA has been done for Japanese: just a few sample analyses were attempted for English.
- MSFA requires, by its very design, an annotator/ analyst to specify a lot of knowledge hard to access for non-native speakers.

Omitted Topics

- MSFA is coupled with a theoretical framework called *Frame-Oriented Concept Analysis of Language* (FOCAL: Kuroda, et al. 2005; Nakamoto, et al. 2005).
 - Competitive Theory of Frame Selection (Kuroda et al. 2006, presented at DGfS) is a product of FOCAL
- But we don't have enough time to talk about FOCAL today.

Overview

- Giving some background
 - Especially why frame definitions and the annotation scheme of *Berkeley FrameNet* (BFN: Fillmore et al. 2003) were not used so far
 - Supplement it with a competitive theory of frame selection
- Presenting sample MSFAs
 - Explain how MSFA goes
 - Try to show what issues BFN will face when full-text analysis/annotation is seriously attempted.
- Summary

How to Annotate Japanese Texts for Semantic Roles with MSFA

Where Does MSFA Come from?

- When Kow Kuroda came to know about the BFN approach to semantic annotation (Johnson and Fillmore 2001; Pinkal et al. 2003) at ACL 2003, he found it really exciting, and wanted to try out the same thing for Japanese text analysis.
- But he faced some difficulties

Major Obstacles

- The following are major obstacles:
 - 1. At that time, the *coverage* of BFN database wasn't broad enough, and as much fine-grained as he needed.
 - 2. What's worse, if we decide to go with BFN frames,
 - A. good understanding of English is required (for both the staff and annotators); this is too selective.
 - B. There will be little chance to link annotation to entries in Japanese thesauri (e.g., *Nihongo Goi-taikei* (A (Comprehensive) Japanese Lexicon)
- While issue 1 is improved greatly in last two years, issue 2 is still a problem.

Decisions Made Two Years Ago

- To annotate Japanese texts for *deep enough* semantics is our objective. So, we decided
 - not to go along with BFN frames
 - to develop our own *scheme* for semantic annotation/ analysis such that
 - it is applicable to given Japanese sentences
 - it provides *deep enough* semantic analysis useful for research in *Cognitive Science of Meaning*
- The following are not our goals *per se*:
 - develop NLP applications like MT, QA system, IR
 - construct a frame database as an "extended lexicon"

Why Semantic Annotation?

- We need annotation/analysis of deep enough semantics to provide the "infrastructure" for *cognitive science of meaning*.
 - Our research is not seriously oriented for NLP tasks, even if there should be no incompatibilities.
- Rather, we addressed the following explorations:
 - Given a sentence *s*, *what kinds of frames* are needed if we wanted to achieve a "psychologically real" description *d* of what (average) people understand when *s* is heard or read?
 - and *how frames* are "put together" in description d?

Far from Trivial Matters

- How to deal with metaphor, metonymy, and other sorts of "figures of speech"
- How to break a sentence into "meaningful units"
 - How to deal with frame-evocation by complex, often discontinuous units?
- How to deal with anaphora?
- How to treat topic marker: is it part of a FE or not?

Short (and Hopefully Gentle) Introduction to FrameNet

What Is a (Semantic) Frame?

• A (semantic) "frame" is

- an organization of "frame elements" (FEs), i.e., situation-specific "semantic roles" in human mind/ brain
- that represents a schematization of situation, or a generalization over "events" (or "states")

• Caveat:

• Don't confuse semantic roles in this sense with socalled "thematic roles," or "deep cases" in the sense of Case Grammar (Fillmore 1968)

Examples

- 1. <Predation>=<<Predator>, <Prey>, ...>
 - i. [Predator A group of killer whales]
 - ii. [Predate.GOVENOR attacked]
 - iii. [Prey a humpback whale].
- 2. <Bank Robbery> = <<Robber>, <Bank>,
 <Weapons>>
 - i. [Robber A group of masked men]
 - ii. [Rob.GOVERNOR attacked]
 - iii. [Target a bank branch in L.A.].

Hierarchies of Frames and Frame Elements

- <Predation> IS-A <Harm-Causation> = <<Harmcauaser>, <Victim>>
 - <Predator> IS-A <Harm-Causer>
 - <Prey> IS-A <Victim>
- <Bank Robbery> IS-A <Robbery> IS-A
 <Attacking> IS-A <Harm-Causation>
 - <Attack> = <Assailant>, <Victim>
- <Attacking> IS-A <Intentionally_act> (IS-A <Agentive_act>)
 - <Bank Robber> IS-A <Robber> IS-A <Agent>

Role-denoting Names

• Note

- "robber," "victim," "predator," and "prey" are all *roledenoting/specifying nouns*.
- But
 - "(a group of) killer whales," "a humpback whale," "(a group of) masked men," and "a bank branch" are not; they are *entity-denoting/specifying nouns*.
- See Kuroda, Nakamoto and Isahara (2006), Gentner (2005), Gentner and Kurtz (2005) for relevant details.

But the Story Isn't Over Yet

- Why is it that sentences 1 and 2 make sense, and 3 and 4 don't, unless they are "adjusted" (metaphorically or metonymically)?
 - 1. A group of *killer whales* attacked a humpback whale. [<Predation> situation]
 - 2. A group of *masked men* attacked a bank branch in L.A.. [<Bank Robbery> situation]
 - 3. A group of *killer whales* attacked a bank branch in L.A.. [<??> situation]
 - 4. A group of *masked men* attacked a humpback whale. [<??> situation]

Samples of "Adjustment" by Semantic Accommodation

- Sentence 3 can mean a <Bank Robbery> if "killer whales" are understood as nicknames for robbers
- Sentence 3 can mean a <Predation> if "masked men" are understood, somehow, to mean a group of <Predator> (e.g., killer whales, sharks,)
- Sentence 4 can mean a <Predation> if "a bank branch" is understood, somehow, to be a nickname for a whale or something of being a <Prey>
- Sentence 4 can mean a <Bank Robbery> if "a humpback whale and her baby" are understood, somehow, to mean a <Bank>

But How Come?

Supplementing Frame Semantics with a Theory of "Frame Selection"



- Each semantic unit SU "activates" a set of frames independently.
- Evoked frames "compete" each other either by mutual "activation" or lateral "inhibition"
- Once competition settles down, the (meaning of) SUs of the "loser" frames "accommodate" to the (meanings of) "winner" frames

Supplementing Frame Semantics with a Theory of "Frame Selection"



- Each semantic unit SU "activates" a set of frames independently.
- Evoked frames "compete" each other either by mutual "activation" or lateral "inhibition"
- Once competition settles down, the (meaning of) SUs of the "loser" frames "accommodate" to the (meanings of) "winner" frames

Supplementing Frame Semantics with a Theory of "Frame Selection"

• Given a sentence $s = w_1 w_2 \cdots w_n$,

- Each word w_i (or (possibly discontinuous) substring
 "... w_i ... w_j ...") "evokes" a frame F_i independently
 each other, and "strengthen" or "suppress" each other.
- Competition among frames evoked takes place, and it "converges" when
 - the specifications of "loser" frames are adjusted to the specifications of the "winner" frames.
- This way, the set $\mathscr{F} = \{F_1, F_2, ..., F_n\}$ of evoked frames reduces into a smaller set of frames \mathscr{F} .

Sample Analysis



Frame Specification Flow of (2)

Feature Specification Flow



Frame Specification Flow of (3)

Feature Specification Flow (Adjustments Required)



Remarks

- Constituency plays virtually no role.
 No effect from [NP [V NP]
- Frame specification flows, with and without bifurcation, should explain the "selectional restrictions" imposed on 1, 2, 3 and 4.
 - The origin of such restrictions are not really lexical one, as suggested by Fillmore in 70's.

Suggestion

- Complex units like the following are able to "evoke" finer-grained, specific frames:
 - "(a group of) killer whales attack ..."
 - "(a group of) masked men attacked ..."
 - "... attacked a bank branch ..."
 - "... attacked a humpback whale ..."
- Typically, this takes place when a role-denoting noun is combined with a verb.
- This fact needs to be considered in annotation tasks. MSFA does it.

MSFA Procedure (Simplified)

- 1. Segment a sentences S into units $U_1, ..., U_n$.
 - This is not independent from Step 2. So, you need to go cyclic.
 - Note incidentally that it's better NOT to try to build up larger units from smaller units. This tends to lead annotators to a "false" analysis.
- 2. For each U_i , find a set of frames $F_1, ..., F_m$ so that one of their "frame elements" is *realized* by U_i .
- 3. Specify relationships among all the frames.

Guiding Principles of MSFA

• "Be meticulous"

- Every word (or morpheme if morphological analysis is necessary) needs to realize at least one "frame element" of a frame.
- You are not allowed to ignore a minor element by saying "its meaning is *uninteresting*." If this "excuse" is allowed, your analysis will get arbitrary very soon.

• "Be greedy"

• To every word, you need to assign as many semantic roles as possible if they are not incompatible

How MSFA Goes –Sample Analysis–

Sample MSFA

- The following is a text taken from *Kyoto University Corpus* (Kurohashi and Nagao 1994):
 - 1. 「ホワイトハウスの内側」という本が十四日,米 国で発売される.
 - 2. 歴代大統領と関係者をこきおろしており、話題になるのは間違いない.
 - ワシントン・ポスト」紙などで長年,調査報道をしてきたロナルド・ケストラー氏の新著.
 例えば次のような内容だ.
 …

Sample MSFA

- The English translations of the text:
 - 1. A book titled "Inside the White House" will go on sale in the U.S. on January 14.
 - 2. The book will definitely be a much-talked-about, severely criticizing the past U.S. Presidents and their aides.
 - 3. The title came as latest work of Ronald Kesler, an expert reporter and investigator at the "Washington Post" and other media.
 - 4. The book, for instance, reveals the following episodes.5. ...

Sample MSFA (for Japanese)

1 4 3	А	В	С	D	E	F	G	Н	1000	1	К	E Care	М	N
1	F-ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
2	F-to-F relations	elaborates F2; constitutes F3		constitutes F5; presumes F5; elaborates F4		presupposes F3	presupposes F5; constitutes F5; presumes F7	presupposes F6; constitues F8; elaborates F9	25/184	presupposes F5	presupposes F9		constitutes F3,F5	
3	Frame idenfitier	作品への題名 つけ	命名	執筆	制作	出版	商品販売	購入	消費	読書	楽しみ	米国での大統 領による為政	秘密の暴露	報告
4	*	1.1.1.1.1.1	19111121			1211-125	1. 1. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2. 2.	1253166	Ch Cherty	13-2722	1.27 8 6.2	The second		報告者
5	*		1922		a per a per fa	目的	1.118	GOVERNOR	GOVERNOR	手段			1942	報告内容 [start1,end]
6	*		a free she to	3422315	111111111	11111	127 Laur	目的	手段	GOVERNOR	手段		239 1743	
7	*	S Sant That	153815 GE	1457683	1212-12-12	P. S. S. F.	12. 10 3213	124.19/2	目的	目的	GOVERNOR	2235466	8-11-51 E	
8	*					小売り業者	販売者	販売者	提供者 [ternary]	t state for				
9	*			122 21 -2	23.22	客	客	購入者	消費者	読者	楽しむ者		A sector	
10	*	題名付与者 [secondary]	命名者[+aux]			出版社			提供者 [secondary]					
11	*	題名付与者 [primary]	命名者	執筆者	作者	協力者?			提供者 [primary]				暴露者	
12	*	S. M. Mart	2.3.5.2.E.	目的1	1.19311-9	2. 2. 2. 1. 1	1943115			and man	1 Partic	6.1.1.1.1	GOVERNOR	
13	Г	MARKER[1,2]	MARKER[1,2]	特徴		出版物.特徴	商品.特徴	商品.特徴	消費物.特徴	読み物.特徴	楽しみ.特徴			報告内容 [start2,end]
14	ホワイトハ ウス	題名	名称									大統領官邸	標的	
15	の												MARKER	
16	内側												秘密: EVOKED	
17	J	MARKER[2,2]	MARKER[2,2]									1. 2 66 14	12 2014	
18	٤	EVOKER1	EVOKER1									114498	697284	
19	題	GOVERNOR	GOVERNOR	EVOKER1									E. M. S.	
20	する	EXTENDER	EXTENDER	EVOKER2										
21	本	作品	命名対象	書物: EVOKER3	作品	出版物	商品	商品	消費物	読み物	楽しみ.源			
22	が	12-5-6-6-64	91 472 F		R. M. C. S. S. S.	MARKER	MARKER	5 3.16 24	1220 8	11 St. 18	and here the	edit all	S. B. S. S. S.	
23	十四	14 2 12 2	S. S. S. S. S. S.	31 - 11 - 18 A	Res Contract	日時:日付	日時:日付	A Starte and		1215282	1. 1. 2. 1. 1.	C. S. Lond Ho	Call and the	
24	B	all a start of the	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	136793	1 - Y - S. S.			129-2112	S. S. Carel		Sof States	VILLES B.		
25	、	145-15-15-15-15-15-15-15-15-15-15-15-15-15		1-2-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1	14 6.34	Sec. 2	12. 36. 72.	122020		1.	12111223		19625 85.24	
26	米国	14 1 1 1 1 1 2		1910-151	4 F. 3 V.E.	場所	場所	516-2352		E122123	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	196524	A. C. S. L. L.	
27	で		501900	5341202	118782	MARKER	MARKER		and a start	1999 420	12 11 20 20			
28	発売		263 5 6 6 7	目的2	Contraction of the	GOVERNOR	GOVERNOR	and the let be	12/22/ 22/2	1. 1. 1. 2. 2.		1. A. A. A. S. A. A.	手段	
29	され	1.22.7. 2	111 11	1.5 12 1. 5 1.	5.30 B.B.B.B.	EXTENDER1	EXTENDER1	Part - Shark		and and	113422	18.18.36	2824220	
30	る	81 × 20 10 10		2000	Station 2	EXTENDER2	EXTENDER2	C. S. S. S. S. S.	CET FARENCE	19-12-19-19		2012 11 12	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	
31	0	Contraction of the		17 () (A) () ()	231 113	-1		CARE AND	1.2.1.2.3	a share a share		11/22/11/2		6 6 62 28

Sample MSFA (for English)

182	А	В	С	D	E	F	G	Н	1	J	К	L	М	N										
1	Frame ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13										
2	F-to-F relations	elaborates F2; constitutes F3		constitutes F5; presumes F5; elaborates F4		presupposes F3	presupposes F4; constitutes F5; presumes F7	presupposes F6; elaborates F9		presupposes F5	presupposes F9		constittues F3,F5											
3	Frame	Title Giving	Name Giving	Writing	Authoring	Publishing	Selling	Purchasing	Consuming	Reading	Having Fun	Presidential Government in the U.S.	Disclosure	Reporting										
4	*	192141	1. a. 2. 1. 3.		and a start of the	2.3472.233			2732224	12 6 20 20	1 1 1 1	19. Jan 5 34	18.2019.20	Reporter										
5	*	-42-54	a la serie a		a start	Purpose	1. 1. 84	GOVERNOR	GOVERNOR	Means	1. States		48.15%	Report[start 1,end]										
6	*	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	1.16 19 28 1	143 Mar 14 Arts	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	11 1 1 1 1 1 1 1	247 6 5 5 6 7 5	Purpose	Means	GOVERNOR	Means	San Start	1214 299											
7	*	Contra Contraction	201222	121 12 12 12 12	1718-59-5-3-1-3	Specific State	1 / K / J / K / K		Purpose	Purpose	GOVERNOR	1201200	6.81522											
8	*	14 10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1	1.3.1.4.1.5.1	1.5. 1. 1. 1.	Retailer	Seller	Seller	Provider3	15 8 64 62		1.2.2.1.1.	12.2.19											
9	*	1	and the second			Customer	Customer	Purchaser	Consumer	Reader	Enjoyer	and the second	11022662											
10	*	Title Giver[seconda ry]	Giver[2]	11-11	Supporter	Publisher	Provider		Provider2															
11	*	Title Giver[primary]	Name Giver[1]	Writer	Author	Supporter?	and	and the	Provider1		a fair fair		Revealer											
12	*	10110		Purpose1	1981 1.55		la se		Sea State	1.201	Domain=Topic		GOVERNOR											
13	A	Work	Object	Book	Work[+Piece]	Publication	Goods	Goods	Commodity	Book	Fun Source	141612		Report[start 2,end]										
14	book												C & 57 78 267 22.	-										
15	titled	GOVERNOR		Book.attribute	Work.attribute	Publication.att ribute	Goods.attribut es	Goods.attribu tes	Commodity.a ttribtute	Book.attribute	Fun Source.attribut e	224												
16	"	MARKER[1,2]	MARKER[1,2]									154120												
17	The	Title	Name									7427	Secrets: EVOKER											
18	Inside											10121252												
19	White																					Presidential Office: EVOKER	Target	
20	House																							
21	"	MARKER[2,2]	MARKER[2,2]									2016 21.61	1.12 20 2000											
22	will	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1			1.2 5.5 4. 2	EXTENDER2	EXTENDER2	18 1 1 1 2 2 4	10/2 2 27 20		12 1 9 24		22-2 EV. A.S.											
23	go on			Purpose2		EXTENDER1 GOVERNOR[+	EXTENDER1 GOVERNOR[+		2 - 5 - 5 - 5 - 5 - 7 - 7 - 5 - 5 - 5 - 5 - 5 - 5 - 5 - 5 - 5	1	1.11.11		Means											
24 25	sale	Contract of		ruiposez	11.397757	composite	composite]		111111				IVIEdITS											
26	in	3982389.24	Sugar in Sat	a spill the start of	15: 1. 8 66 2. 3.	MARKER	MARKER	220 6781455	15. 2. 19 1 and	S/ S	1122346753	2525 8 222	2020200000											
27	the	Sarah Salar	1012200	1.1919-1.29-53	8-15 5. 5. 7. 5	Place	Place	13.4.252.2	C. S. S. S. A.	- F1 2223	13512619	5 12 65 20	12.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.2.											
28	U.S.	1.5516.221	13 5 1 7 93	19.5.5 2	121111111111111			Class and	and state of all	16 32 19 24	1000000	CALL STORAGE AND	18222400											
29	on	222 1 1 2 2 5 7	11 27 S. 17	2233 800 80	Saw Press	MARKER	MARKER	11. 16 S. 5 . +	Ser Strate	1.15 Sciel 1	5000 401	1 1 - 1 1 2	12000											
30	January	11 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	and a tribe	1949 6 43	23.457633	Time: Date	Time: Date		The Martine States	13347511	State Carta	1 1 1 1 1 1 1 1	220220											
31	14		S. 1.	1. 1. 1. 1. 1. 1. 1.				ast a standay	47.5.16.7.25	de la composition	74442	14 19 10 20	12-20-27											
32		203 + 250		E. C. C. C. C. C. C.	and all place	and start		See Anter 1 1	1 2000 1000	145 S & T. C. C.		2293629993	100000000000000000000000000000000000000											

Sample MSFA (for Japanese)

$\begin{array}{c c} A \\ A \\ F - IC \\ F - to-relation relation Fram idenfit 4 * * * * * * * * * $	ID O-F tions Imme tifitier * * * * * * * * * * * * * * * * * * *	B F1 elaborates F2; constitutes F3 Title Giving Title Giving Giver[seconda ry] Title	ary]	D F3 constitutes F5; elaborates F4 Writing	E F4 Authoring	Publishing Purpose	G F6 presupposes F5; constitutes F5; presumes F7 Selling	H F7 Presupposes F6; constitues F8; elaborates F9 Purchasing GOVERNOR	F8 Consuming	F9 presupposes F5 Reading	K F10 presupposes F9 Having Fun	F11 Presidential Government in the U.S.	M F12 constitutes F3,F5 Disclosure	N F13 Reporting Reporter
F-to-relation 2 3 Fram 3 6 7 8 9 1 2 3 1 2 3 1 2 3 7	xo-F tions	elaborates F2; constitutes F3 Title Giving Title Giving Giver[seconda ry]	Name Giving	constitutes F5; presumes F5; elaborates F4		presupposes F3 Publishing Purpose	presupposes F5; constitutes F5; presumes F7	presupposes F6; constitues F8; elaborates F9 Purchasing GOVERNOR	Consuming	presupposes F5 Reading	presupposes F9	Presidential Government	constitutes F3,F5 Disclosure	Reporting
3 idenfit 4 * 5 * 6 * 7 * 8 * 9 * 0 * 1 * 2 * 3 f * 7 * 8 * 9 * 1 * 1 * 1 * 1 * 1 * 1 * 1 * 1	ifitier * * * * * * * * *	Title Giver[seconda ry] Title	Name Giver[second ary]	Writing	Authoring	Purpose	Selling	GOVERNOR			Having Fun	Government		
5 6 7 7 8 8 9 * 9 * 0 * 1 2 * 3 1 * 7 * * * * * * * * * * * * *	* * * * * * * * * * * *	Giver[seconda ry] Title	Giver[second ary]			847 (78) 862277			GOVERNOR	1000000074		S. S. S. S. S. S.	and the second	Reporter
5 6 7 7 8 8 9 * 9 * 0 * 1 2 * 3 1 * * * * * * * * * * * * *	* · · · · · · · · · · · · · · · · · · ·	Giver[seconda ry] Title	Giver[second ary]			847 (78) 862277			COVERNOR					
6 * 7 * 8 * 9 * 0 * 1 * 2 * 3 「 4 ウス	* * * * * * * *	Giver[seconda ry] Title	Giver[second ary]			Dete			GOVERNOR	Means	S. S. S. A.S.	Contraction of the		Report[start1 ,end]
8 9 4 0 1 2 3 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	* *	Giver[seconda ry] Title	Giver[second ary]			Dete:	Take State	Purpose	Means	GOVERNOR	Means		12 22 20 20 10	,enu]
8 9 () 1 2 3 () * 2 3 () * 3 () * 4 2 * 1 * 1 2 * * 1 * * * * * * * * * * *	*	Giver[seconda ry] Title	Giver[second ary]			Deteller	State and the state of the state	- arpose	Purpose	Purpose	GOVERNOR		5. <u>199</u> - 11 - 10	
9 1 2 3 5 7 7 7 7 7 7 7 7 7 7 7 7 7	*	Giver[seconda ry] Title	Giver[second ary]			Retailer	Seller	Provider	Provider[tern		12142			
0 * 1 2 * 3 3 「 4 ウス	*	Giver[seconda ry] Title	Giver[second ary]	3. C. S.		Customer	Customer	Purchaser	ary] Cosumer	Reader	Enjoyer			
1 2 ・ * 3 「 4 ウス		Title				Publisher	Customer	Turchaser	Provieder[sec ondary]	Reduct	Lijoyei			
3 「 4 ホワイ クス	*	Giver[primary]	Name Giver[primary]	Writer	Author	Supporter?			Provider[pri mary]				Revealer	
3 ホワイ 4 ウス				Purpose1						1.28.20	ald all	1.2349-63	GOVERNOR	
4 ウス	Г	MARKER[1,2]	MARKER[1,2]	Book.attribu te	Work.attrib ute	Publication.a ttribute	Goods.attribu te	Goods.attribu te	Commodity. attribute	Book.attribut e	Fun.attribute	1111		Report[start2 ,end]
5 の		Title	Name									Presidential Office: EVOKER	Target	
	の												MARKER	
6 内側	側												Secrets: EVOKER	
7		MARKER[2,2]	MARKER[2.2]									25 1265		
8 2		EVOKER1	EVOKER1									548967	11-1-1-1-1-	
9 題		GOVERNOR	GOVERNOR		EVOKER1							222226		
20 する		EXTENDER	EXTENDER		EVOKER2							S. S. S. S. S.	1000000000	
21 本		A Piece of Work	Object	Book (as a Piece of Work)	Work	Publicationk	Goods	Goods	Commodity	Book (as Information Carrier)	Fun Source			
22 が	<u> 5</u> 5	1. 6 . C. C. C. C.	1125.65	137300		MARKER	MARKER	12 5 5 1 2 3	1. 20 C .	1995 6	The start	S. P. S. S.		
23 + 🗠		1 1 1 2 1 (ST		S. P. C. P. S. St. P.	S. C. S. S. S. S.	Time: Date	Time: Date	- Restricted		S. S. S. S. S.	SPI CALLED	grad and	1415-20 M	
24 🛛 🖪			1 1 2 2 2 4	136294	Start and the			12.24 279	A Start	A.S. Salata	The second second	Vinte Jon Star	1000	
25		and the second		The share of the	14 19 11	22312 142		1.	12:41 27	C.S.A.S.	12 12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		1 4 4 6 A 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
26 米国	(国	1.1.5-9.8.2	Para de la la	200102	1. 5 25 1	Place	Place	121220	A State States	S. 20022	A PAR PAR	State Provide		
27 で		1257000000	126135	1. 1. 1. 1. 1.	11111111	MARKER	MARKER	541444	191823	123119	128 3.58	1. 1. 2. 2. 1.	12122	
28 発売		S. Strates	20 2 2 2 2 2 3 3	Purpose2	1.6 1	GOVERNOR	GOVERNOR	1999 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	6 99 92 F 19	Call A St.	all for a	1. 6. 5. 2. 1.	Means	
29 され		San Star		1000000000	5816010	EXTENDER1	EXTENDER1	State State	19.05-2.60	and the first	17919999	NGE DEN	1992022	
30 3	<i>X</i> .	100000000	1.2.1.8.1.3	19 48 19 19		EXTENDER2	EXTENDER2	and a start	Ser and	Sec. 19820	Sel Cont		and the second second	
31				2001 2000	2012250	1. 1. 2. 1. 1. 1.		A CONTRACTOR OF A						The same of the second second

Remarks

• MSFA is used:

- to identify and specify as many frames as possible; each column, with a "Frame ID" (local variable) and a "Frame Name" (global variable) specifies a frame
- to specify explicitly how frames are interrelated using Frame-to-Frame Relations (global) on the second row

Conventions

- The relative order of columns is not significant.
- "Null instantiations" are indicated by * if they are position-neutral, and by ** if they are position-specific

What MSFA Does?

• MSFA above specifies, for example:

- *Author,* as a role-denoting noun, designates an Agent-class role specific to <Authoring>.
- Writer, as a role-denoting noun, designates an Agentclass role specific to <Writing>, a subclass of <Authoring>

• etc

• This implicitly describes frame hierarchies and role/FE hierarchies like ...


Frame-to-Frame Relations

- The partial list of Frame-to-Frame relations we have defined so far is:
- "F elaborates G" (deals with Inheritance, "Is-A")
- "F constitutes G" (deals with "Part-Of" relation)
- "*F* presupposes *G*", "*F* negates *G*" (deals with "implications")
- "F motivates G" (can be used to specify <Reason>)
- "F realizes G" (can be used to specify <Purpose>)

Benefits of Multilayered Analysis

- Multilayered analysis has its own benefits.
 - it allows us to explore the details of *frame-to-frame relations*
 - Full text analysis tells much more about them
 - it allows us to explore and specify *multiple*, *simultaneous role realization* by a lexical material.
 - it allows us to avoid *frame conflation*
 - this happens all the time
 - by disentangling a complicated relationship among frames evoked in a sentence

Role Multiplex and Relativized Role

- In the following sentence, the role of 松葉づえ(で) is essentially ambiguous:

 (1) 研究室で友人と話していると、そこに太郎が松 葉づえで入ってきた。
- in that it realizes at least the following two roles simultaneous and inclusively:
 - R1: <歩行者>「太郎」にとっての<歩行の道具> IS-A <Instrument>
 - R2: 観察者 (語り手とその友人にとっての<出現の
 際の様態> IS-A <Manner>

例文(1)のMSFA

-	1.1000	1012		an and a	al al al al	1.	22 22 2	1000	1.1.2.2.	2.27.22	and the	MY RIVER		10000			1.000	Carlos and	22 / 24av			1	146.28
Frame ID (Local)	F01	F02	F04	F03	G05a	G05b	F06	F07	F09	F10	F08	F11	F11	G01	G03	F14	G02	G07	F12	G06	F13	F15	F16
Frame-to- Frame Relations (Global)		constitutes F04	consititute s F01	constitutes F04; integrates G05a,G05b			presuppose s F07; constitutes F09		presumes F08; elaborates F10; constitutes G05a	elaborates G05a		constitutes F14,G02	consititute s G01,G03	constitutes F03; presuppose s F14,G05b; results_fro m G02		elaborates G02,G05b, G07	?elaborate s G07		presuppose s F13; realizes F14; elaborates G06	not_presup poses F13; realizes F14		presuppose s F16; constitutes G05b	presuppose s F13; constitutes F12
Frame Name (Global)	語り	~指定[過去 性の]~	~指定[事態 の前後関係 の]~	切り替わり [場面の]	事態	事態	研究	?説明[用途 の]	会話	維持[状態 の]	交友	~参照~[1]	~参照~[2]	出現[場面へ の]	観察[事態 の]	入室	入場[一般的 な意味での]	位置移動	歩行[足にケ ガをした人 の]	歩行	負傷	発見[不自由 性の]	患い[不自由 性の]
*	聞き手			Stall 4				12.02.03		712.2	1200		12 294	1.2.1.1.1	11241	Ser Charles		R. 1. 2. 79		1.0.723	147,92		1921 k 1/2
*	語り手	指定者	指定者	場面[1]	参与者[1]		研究者 [potential] [1]	利用者[1]	参加者[1]	維持者[1]	交友者[1]	参照者	参照者	観察者[1]	観察者[1]					See.		知覚者[1]	
研究	内容	事態	事態[1]		場所		GOVERNOR	目的[利用 の]	場所	場所	場所	参照先	参照先	観察者 [1,2].Attr[1,2]	観察者 [1,2].Attr[1,2]								
室						S. A.S.	場所	GOVERNOR									1.1.4.1	Call of the Call	58647		Sec. 2		1500
で					MARKER	- 65 - 5			MARKER	MARKER	MARKER	1244						Children 15					- 1
友人					参与者[2]		研究者 [potential] [2]	利用者[2]	参加者[2]	状態	相手: EVOKER			観察者[2]	観察者[2]							知覚者[1]	1
لح					MARKER				MARKER		MARKER			観察者 [1,2].Attr[2,2]	観察者 [1,2].Attr[2,2]								
話し					参与の仕方	912.29	15120	222	GOVERNOR		?手段	2294				1000 S.A.	2.200 2	181153			63999	42000	
て						1000	1225	1.5	状態	GOVERNOR	MARKER					21.72	125-2-1	1226	1.25.25		12 - 2 - 2		133494
いる と			EVOKER	EVOKER	101222		5	200217							22.557.5			34447 - SA	C		1.2.0		
			EXTEDNER	EXTDNER	1.19	277 10 10	11.000		012.9			-	199108 11			1. g. g.m. 15		1111		2100	142 / Sa	3.4.1.0	1.813
そこ			事態[2]	場面[2]	4462	場所	14400	12 2 4 A.	4280	3/1844 8	1912	参照元	参照元	場面	場所	目的地点	場所	目的地点	到着地点	到着地点	a1.3	3359	500000
10					1296	MARKER	17.2 8.8	6.11.19	1840	2 11 1	18042 M		1.2.2.	MARKER	and and here	MARKER	MARKER	MARKER	MARKER	MARKER	Provinsi		1.112.00
太郎						参与者	研究者 [potential] [3]	利用者[3]			相手 [potentila]			出現者	事態	入室者	入場者	位置移動者	步行者	步行者	負傷者	不自由者	?患者
が					1945.252	MARKER	153.20	S. 87 6.31	- 1. A.	2111			S. 6. 5 5 M.	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER
松葉づえ						参与の仕方								様態	様態	様態[入室 の] OR 手段 [入室の]	様態[入場 の] OR 手段 [入場の]	様態[入場 の] OR 手段 [入場の]	道具		EVOKER	様態: EVOKER	様態: EVOKER
で					12.2.2			1816	12000	68113	24350		12.84	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	10-10-62			
λο					1997				1748		24	- 101		EVOKER	EVOKER		GOVERNOR	EVOKER	EVOKER	EVOKER		不自由性を 伴う行動	不自由性を 伴う行動
てき									1 2 2 3				14.14	-		EXTENDER	EXTENDER	EXTENDER					
さた		GOVERNOR				EXTEDNER		1. 1. 1. 1.	1	1 / 1	1.4.7												
	EXTENDER		EXTENDER	EXTEDNER	122	37	0282		1.5.00		11 4		1.1.1.1	11.5.18	18 3 1 (BY)	1222	P. L. R. S.	1-22			28.2.5%		11.750
1	a subscription of the	600 No. 19. 19.	and the second second	1. 2. 2. 1. 2. 2.	1 2 ST. 1 CT. 1 CT.	1. ST46 21.	10 C 10 C 10 C	1.1.2.2.2.4	10000	1	2 1 1 2 2 A 1 3	6-2 4-17	Contraction of the	Section of the	a the standing	166 22	and the second	and the second second		Carrow Marson		and the stand of the	Constanting of the

<u>http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/hiki.cgi?c=view&p=msfa-</u> <u>matsubadzue-DE</u>

Network of Frames



http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/hiki.cgi?c=view&p=msfamatsubadzue-DE

	出現[場面へ の]	観察[事態 の]	入室	入場[一般的 な意味での]	位置移動	歩行[足にケ ガをした人 の]	歩行	負傷	発見[不自由 性の]	患い[不自由 性の]
*		4. 18 M (1 - 1)			1.21.92		Call In		States.	
*	観察者[1]	観察者[1]							知覚者[1]	
研究	<mark>観察者</mark> [1,2].Attr[1,2]	観察者 [1,2].Attr[1,2]								
室		, -	and the second				14 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			N. Later
で					and the set					
友人	観察者[2]	観察者[2]							知覚者[1]	
٤	観察者 [1,2].Attr[2,2]	観察者 [1,2].Attr[2,2]								
話し		, ,	14239187		4.649.8	24 445	14 19 19 19 19		Berghe	april and
τ					alt free		1999 A.S.			
いる			1000	1 Land			1.1.1.2.5		12. 19	12933
٤				111111		and the set			14114414	
		18-2							12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
そこ	場面	場所		場所		到着地点	到着地点	215 61 8 JA	Start Contractor	
<u>に</u> 太郎	MARKER 出現者	事態	MARKER 入室者	MARKER 入場者	MARKER 位置移動者	MARKER 步行者	MARKER 歩行者	負傷者	不自由者	?患者
が	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER
松葉づえ	様態	様態	様態[入室 の] OR 手段 [入室の]	様態[入場 の] OR 手段 [入場の]	様態[入場 の] OR 手段 [入場の]	道具		EVOKER	様態: EVOKER	様態: EVOKER
で	MARKER	MARKER	MARKER	MARKER	MARKER	MARKER	6.11.20			
ک ې	EVOKER	EVOKER		GOVERNOR	EVOKER	EVOKER	EVOKER		不自由性を 伴う行動	不自由性を 伴う行動
て			EXTENDER	EXTENDER	EXTENDER					
き										
た				and a start of the second second						
o										

Web Sites that Host MSFA Samples

• FOCAL Hiki Site 1

- http://www.kotonoba.net/~mutiyama/cgi-bin/hiki/ hiki.cgi?FrontPage
- Unrestricted access to MSFA-based sample annotation of copyright-free J-E translations

• FOCAL Hiki Site 2

- http://www.kotonoba.net/~mutiyama/cgi-bin/hiki2/ hiki.cgi?FrontPage
- Restricted access (you need an account) to MSFAbased sample annotation of J-E newspaper article alignment data.

Dealing with Multiplicity

WSD Needs to Be Done Frame-wise

- A single "entity" referred to in a text is likely to *realize different frame elements frame-wise simultaneously*.
- For instance, *book* in the previous example realizes:
 - <Information Carrier> FE in <Reading> frame
 - <Goods> FE in <Selling/Buying> frame
 - <Piece of Work> FE in <Writing> frame
 - <Publication> FE in <Publishing> frame
- This means that word sense disambiguation *needs to be done frame-wise,* explaining why WSD isn't enough for text understanding.

How to Deal with "Markers"?

- Multiple, simultaneous realization of a set of FEs poses the following problem:
 - How to treat prepositions in languages like English and postpositions in languages like Japanese?
- This boils down to the following question:
 - Are prepositions or particles really *parts* of FEs, or just *markers* of them?
- We adopted the idea that markers can be (and need to be) separated from FEs themselves to facilitate encoding of multiple role realization.

For Future

Current Status

• So far, MSFA was applied to

- a portion of copyright-free Japanese-English alignment data (67 sentences), open data
- a tiny portion of *Kyoto University Corpus* (KUC) texts (3 articles, 63 sentences); semi-closed data

Characteristics

- MSFA assigns a sentence 20-30 frames on average (depending on the length)
- 2157 frames (=927+1227: type count), 6846 frame elements (=3031+3815: type count) are recognized
- But no evaluation is done yet.

What Should We Do Next?

- Our own frame definitions need to be linked to BFN ones. But, at the time being, it seems to be unrealistic to establish a mapping between BFN frames and the frames we defined with MSFA
 - granularity difference matters
- But it would be useful if we try to *share an annotation scheme* to facilitate such mappings in the future, and we want to try out best to do it.
- Putting aside the granularity issue, the major problem is with the treatment of markers like prepositions and postpositions.

Thank You

Acknowledgements

My research is indebted to contribution by the following people:

Toshi-yuki KANAMRU, Kyoto University Keiko NAKAMOTO, Bunkyo University Jea-Ho LEE, Hajime NOZAWA, Yoshikata SHIBUYA, Masao UTIYAMA, NICT Koichiro NAKAMOTO