

沖縄語

沖縄語 (Okinawan / うちなーぐち) とは・・・日琉語族琉球語派北琉球語群に属する言語。UNESCOに危機に瀕した言語として認定されている (Central Okinawan)。



Universal Dependencies

ツリーバンクの開発

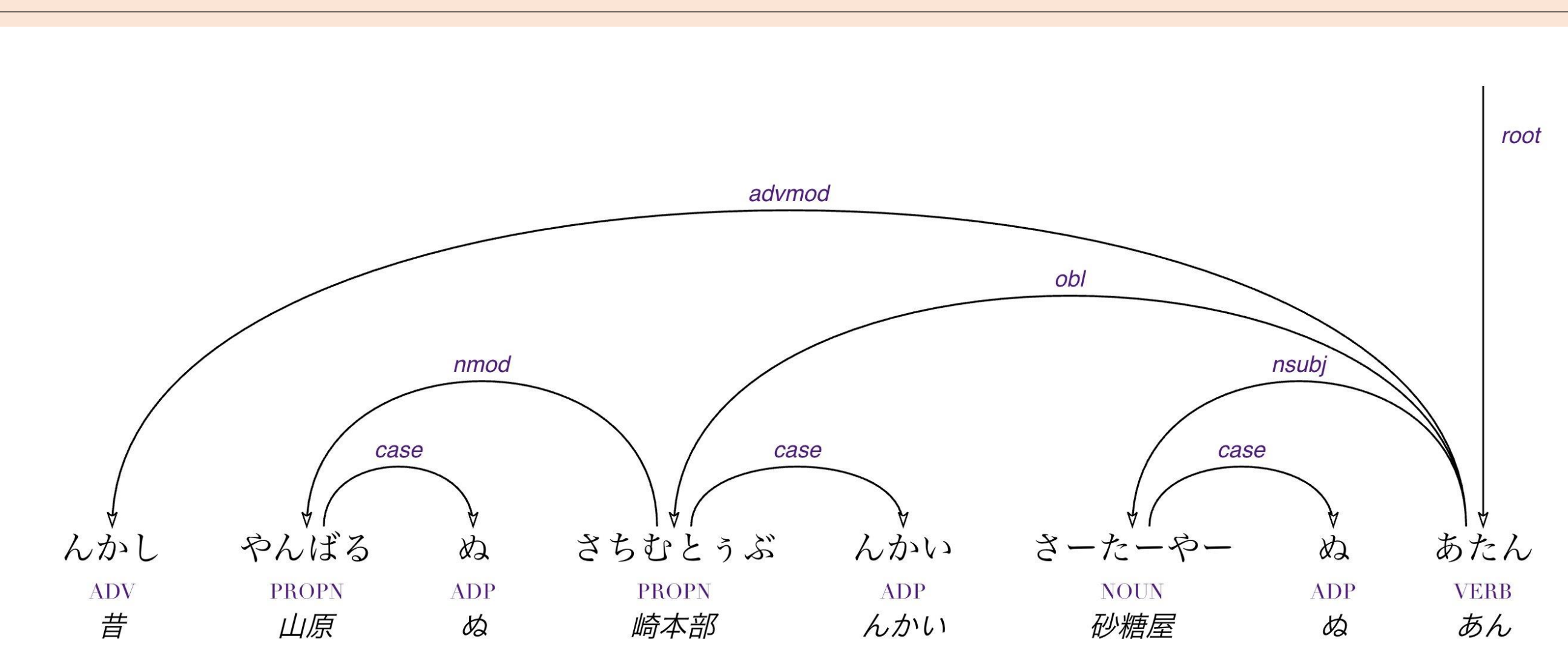
京都大学

宮川 創

沖縄語ツリーバンクの教師データを作成中→完成したらTransformerとSuParを使ったesuparで機械学習

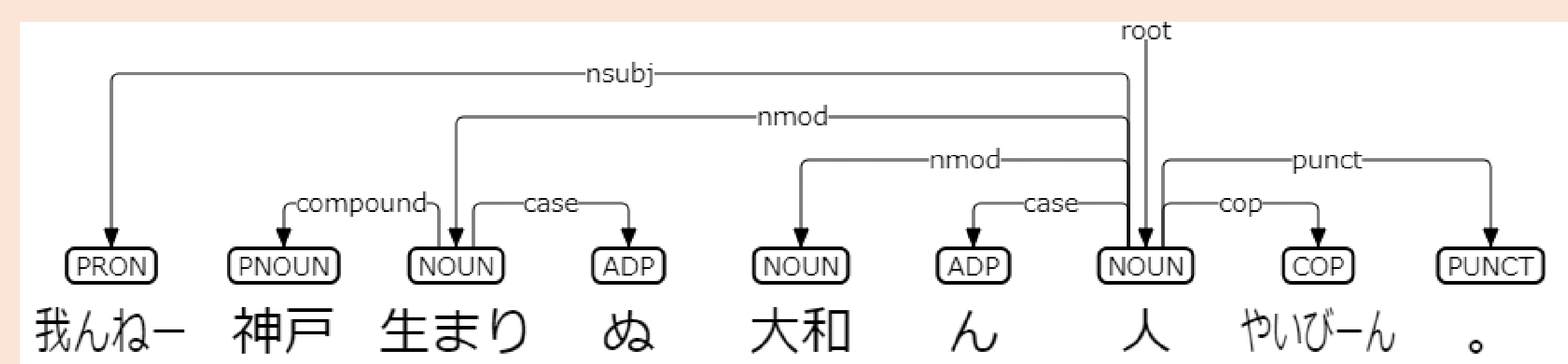
スピンオフとして誰でも使いやすい辞書アプリを構想中

様々な視覚化が簡単・効率的に



Arborator Grew

esupar (解析) x deplacy (表示・編集)



全ての言語で同じ方式でコーパスを

どの言語でも **変わらない** 17の普遍的品詞タグ (UPOS) と 37の統語関係タグ

Input (editable):

1	んかし	昔	ADV	ADV	-	8	advmod	-	-
2	やんばる	山原	PROPN	PROPN	-	4	nmod	-	-
3	ぬ	ぬ	ADP	ADP	-	2	case	-	-
4	さちむとうぶ	崎本部	PROPN	PROPN	-	8	obl	-	-
5	んかい	んかい	ADP	ADP	-	4	case	-	-
6	さーたーやー	砂糖屋	NOUN	NOUN	-	8	nsubj	-	-
7	ぬ	ぬ	ADP	ADP	-	6	case	-	-
8	あたん	あん	VERB	VERB	-	0	root	-	-

普遍的な品詞タグと統語関係タグ

沖縄語の文法的特性が計量的にわかる

例えば、122言語で語順の割合が**統計的**にわかる (SOV, SVO, VSOなど...)

通言語的・通時的な統語構造の計量化
言語記述・言語類型論への貢献

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	ID	辞書 ページ	見出し語	漢字かな現代	漢字か ひらがな	カタカナアクセ ント	UPOS	品詞	文語な どの 種別	補足	意味 1.	意味 2.	意味 3.		
1															
130	129	103	?agai		あがい	①	NOUN	名			地面の高いところ。かみ。sagaiの対。				
131	130	103	?agaihana	上がいい端	あがいはな	①	NOUN	名			日の上りはじめ。				
132	131	103	?agaitida	上が太陽	上太陽 あがていーだ	①	NOUN	名			上る日。朝日。～du'uganuru,sagaitidaa'ugamaN. 上る日は拝むが、落ちる日は拝まない。権勢のよい者につく意。				
133	132	103	?aga=juN	上がゆん	あがゆん	①	VERB	自=raN=ti			1.上がる。昇る。空間を上げる。また。物価などが上がる。tidanu～.日が昇る。maainu～.まりが上がる。niinu～.値が上がる。nubujuNの項参照。	2.上達する。tiinu～.3.できあがる。ま			
134	133	103	?agami=juN	崇みゆん	あがみゆん	①	VERB	他=raN=ti			あがめる。敬う。				
135	134	103	?agane		あがねー	①	NOUN	名			儉約。節約。経済。～nu'jutasjaN.節約がうまい。				
136	135	103	?agane=juN		あがねーゆん	①	VERB	他=taN=ti			節約する。儉約して大事に使う。mookirazijaka ?aganeeri.もうけるよりも儉約せよ。(ことわざ。mookiraziは形は否定だが意味は肯定。mookijuSijakaともいえる。)				
137	136	103	?agari	東	あがり	①	NOUN	名			東。?ni(西)の対。				
138	137	103	?agari	東江	あがりえ	①	NOUN	名			東江。《地》参照。				