

The Areal & Genealogical Dimensions of Hiatus

Kristine A. Hildebrandt, The University of Manchester

kristine.hildebrandt@manchester.ac.uk

I. What is Hiatus & Why Consider it Here?¹

- Two vowels become adjacent via morphological (or syntactic) concatenation:

(1) Persian (Indo-European) (Mahootian 1997)

mi-xa-æm > [mi.xam] DUR-want=1SG 'I want'

- Across languages, hiatus environments subject to different 'responses':

(2) Possible Responses

"NOTHING"

a. Heterosyllabification

Digueño (Yuman)

nya-i:ma > [nja.i:ma] WHEN-he.dances (Langdon 1970)

"RESOLUTIONS"

b. Vowel elision/deletion (Note: in different lgs, different vowels may delete under different circumstances)

Cayuvava (Isolate)

ki-učārāhi > [kučārāhi] SUB-he.is.coming

c. Consonant insertion (Note: in different lgs, different consonants under different circumstances)

Kusunda (Isolate)

tui-i > [twi.ji] bee-GEN (Watters 2005)

d. Glide formation

Luganda (Benue-Congo)

ki-aa > [kjaa] 3P-GEN.LKR (Hyman & Katamba 2003)

e. Morpheme coalescence/merger

Dolakha Newar (Sino-Tibetan)

ye-u > (jeu)σ come-NR1 (Genetti 1994)

f. Allomorphy

Nepali (Indo-European)

Palpa-ilo 'Palpa-ADJ' > [palpali] 'like Palpa' (a town) (Acharya 1991)

- Autosegmental approaches (e.g. Sanskrit & Chicano Spanish; Schane 1987)
- Domain-oriented approaches (e.g. Isoko; Donwa-Ifode 1985)
- Optimality Theory & alignment constraints (e.g. vowel elision in Casali 1995; 1997)
 - Positionally Sensitive Faithfulness:
 - MAXLEX: maintain phonological material in elements encoding greater semantic content (root/lexical morphemes).
 - MAXWD: maintain phonological material in word-initial position

¹ The data & inspiration for this work come from the Word Domains component of the Autotyp Project (www.uni-leipzig.de/~autotyp). Thanks to Franziska Crell and Thomas Goldammer for helping me collect data from Indo-European & Austro-Asiatic languages, in particular, but any errors are my own responsibility. Ultimately, my representation of phenomena here is dependent on exhaustivity/comprehensiveness of grammatical descriptions & feedback from researchers. If you see a case here that is mis-represented or if something is omitted, feedback is welcome!

II. Goals & Methods

A. Explore 'hiatus' in cross-linguistic perspective

- Findings mixed significance
- Exploratory endeavour & matching of observations with other, related findings in the Word-Domains project
- Critical appraisal of past analyses in languages with 'mixed' responses & resolution types

B. Language Sample

- 45 languages (Appendix)
- "Balanced" representation of 3 families:
 - Indo-European, Sino-Tibetan, Austro-Asiatic, plus a mixture of 'other'
- "Balanced" representation of 3 areas:
 - Europe, Indic/S. Asia, Southeast Asia, plus 'other' areas
- Final sub-sample: Indosphere, Buffer Zone & Sinosphere sub-areas for Sino-Tibetan only
- All languages coded for:
 - Whether a hiatus environment is possible (i.e. is syllable canon only min. /CV/?)
 - Whether or not, if an environment is possible, there is a resolution or heterosyllabification
 - Resolution type
 - The morphological domain (i.e. stem-suffix; prefix-stem, affix-affix, etc.)
 - What is affected (e.g. if deletion in stem-suffix domain, what is deleted)
 - Multiple different resolutions in a single domain/edge or 'mixed' resolutions (i.e. a mixture of heterosyllabification and resolutions in same domain)

III. Selected Findings

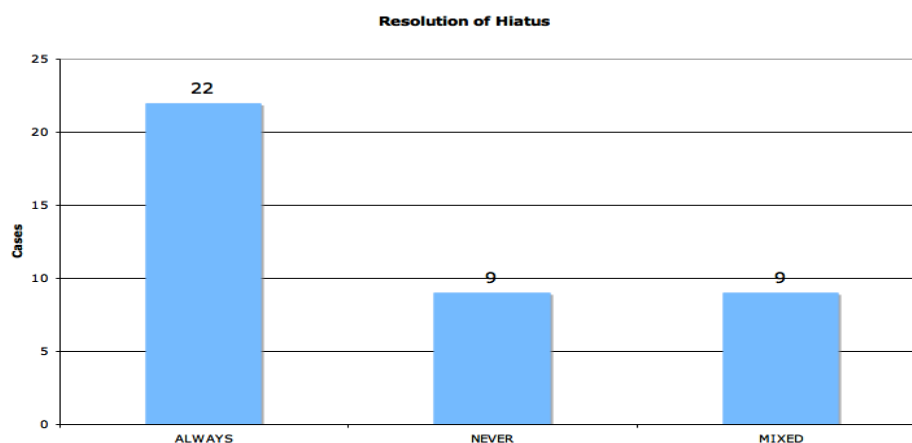
A. Resolutions: Presence/Absence/Mixed

- Languages with minimal syllable template of /CV/ excluded

Question 1: Of languages that display hiatus environments, how often is it "resolved"?

- From OT Perspective, this is not really a question to be asked
- From "empirical" perspective (e.g. Blevins 1995; Maddieson 2005), [CV] a preferable structure

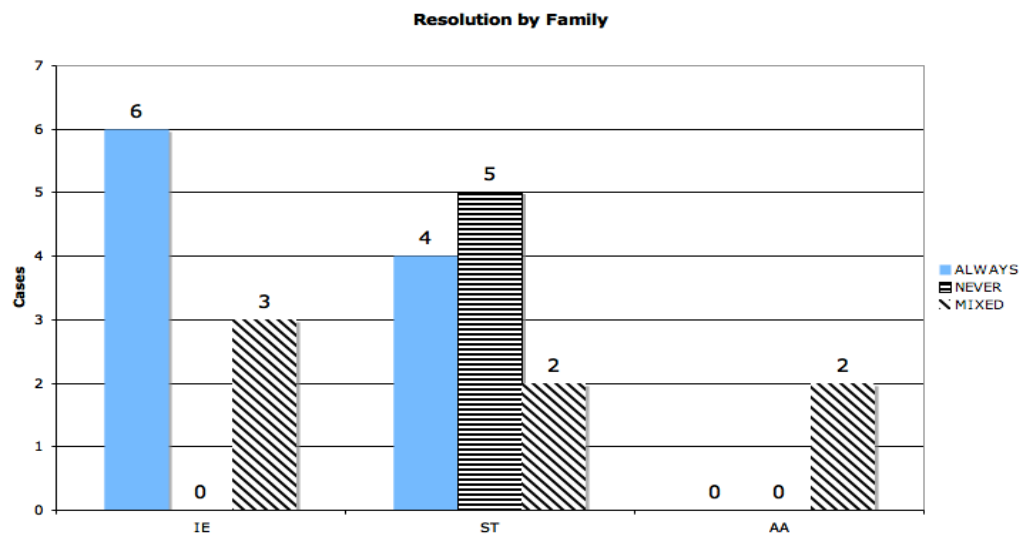
Fig. 1. Response across all languages (n = 40)



Answer 1: If we do not expect a preference for [CV], then resolution is a significantly preferred response to hiatus across all languages in the database (Chi-Square; $p = .017$)

Question 2: Is this preference similar in languages of the three sampled families?

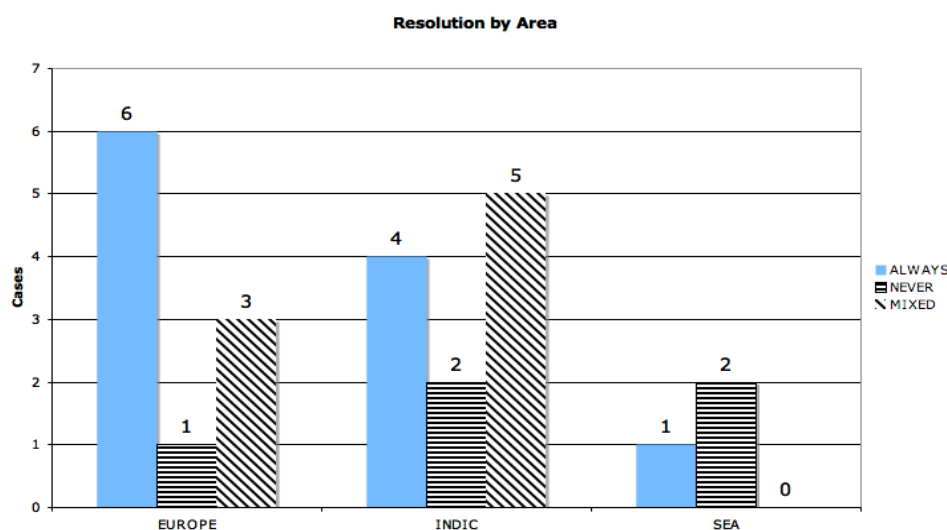
Fig. 2 Response across three families



Answer 2: In Indo-European, resolution is preferred; in Sino-Tibetan all responses are attested; in Austro-Asiatic, languages only have mixed responses; These patterns significantly differentiate the 3 families. (Chi-Square; $p = .040$)

Question 3: Is this preference similar in languages of the three sampled areas?

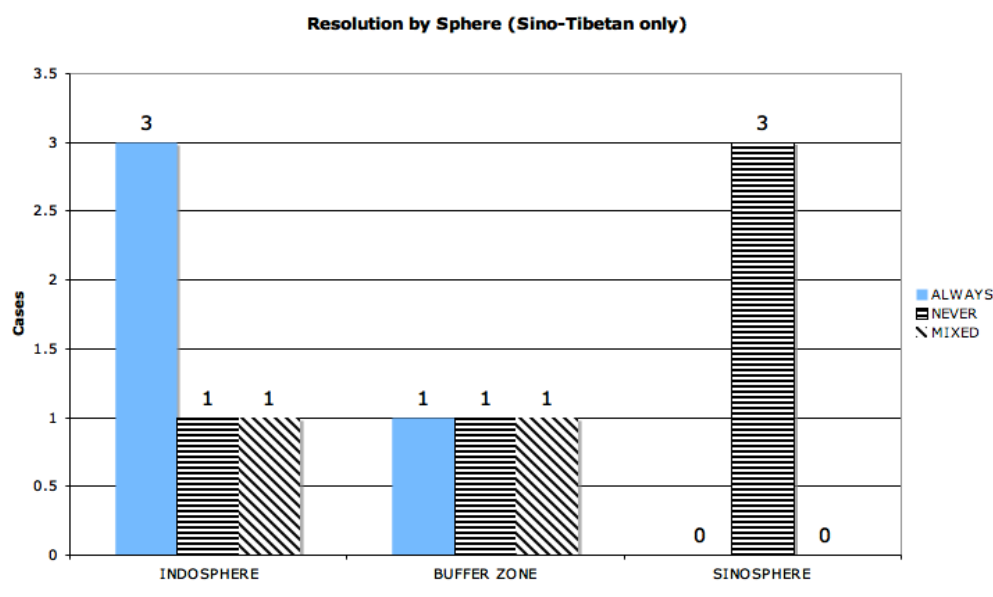
Fig. 3. Response across three areas



Answer 3: In Europe & Indic all responses are attested; In SEA languages either always resolve hiatus or never resolve it (heterosyllabification). These patterns do not significantly differentiate the 3 areas.

Question 4: Is this preference similar in the languages of the three (proposed) sub-areas for S-T?

Fig. 4. Response across three spheres



Answer 4: In Indospheric & Buffer Zone languages all responses are attested; in Sinospheric languages heterosyllabification is the only response. These patterns do not significantly differentiate the 3 spheres.

- The 'global' trend is resolution (all things assumed equal) & this trend is supported in Indo-European languages
- In the three areas and three spheres responses are more mixed
- Interesting to note that Sinospheric languages permit ONLY heterosyllabification

Question 5: Given the cross-linguistic tendency for hiatus resolution, can we expect languages without resolution to have an overall lower "coherence" profile?

- Coherence (COH) (Bickel & Hildebrandt 2005): A ratio of a language's available morphology to number of these morphemes included in phonological words in that language.

(3) Belhare (Sino-Tibetan; data from Bickel & Hildebrandt 2005)

- Available Morphology: root, prefix, suffix, circumfix, enclitic
- Big P-word: 4 of 5 (Final Velar Drop)

a. ka-ak-lu-kak=phu>(kaaklugakphu)_ω 1sP-OPT-tell-2A=REP 'You may tell me, they say.'

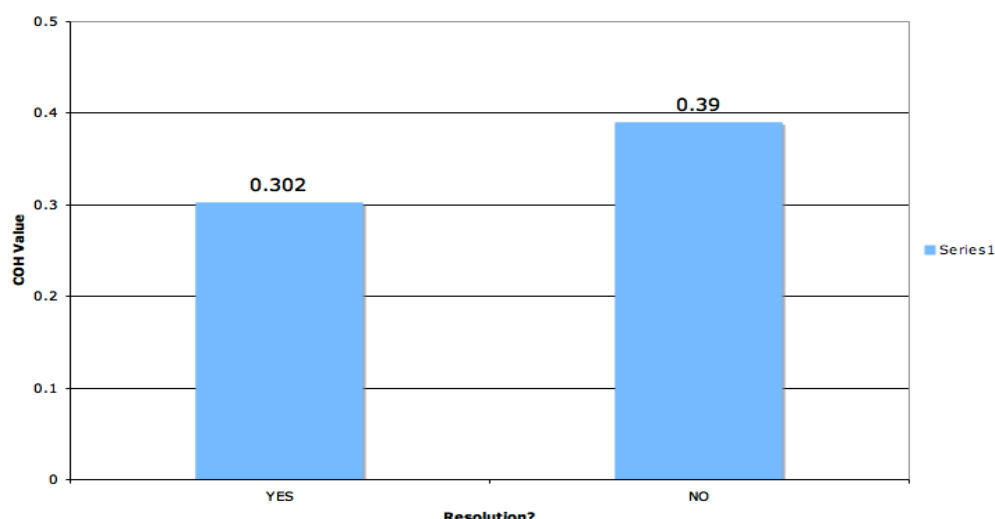
Small P-word: 2 of 5 (Nasal P.O.A. Assimilation)

b. mi-N-kai-chi-n > (miŋkai)_ωchin '3nsS-NEG-come.up-d-NEG'

Overall COH value for Belhare is .25 (about middle so far)

- A typological variable based on exhaustive surveys of phonological word types in a similar sample of languages (Austro-Asiatic languages excluded here due to lack of COH values)
- Smaller COH value (approaching 0) = smaller P-words; Bigger value = bigger P-words

Fig. 5 Average COH values



Answer 5: A significant ($p = .045$) correlation between the PRESENCE of hiatus resolution and SMALLER COH values (& between the LACK of resolution and LARGER COH values). Suggests that we may not be able to rely on a particular process to indicate the overall coherence profile of a language. Thus, certain, highly salient processes may give a false impression of a language's COH profile. This trend is repeated in the family, area and sphere samples, but is not significant.

B. "Conflicted" Hiatus

- Two types

i. Different Resolutions, Same Domain

- Languages resolve hiatus, but we see unique resolution types happening in the exact same domain/edge
- Suggests that lexically specified details are necessary to describe the processes

(4) Persian

different resolution types with æm 1.SG pronominal enclitic, depending on domain type

a. $\text{ketab-ha}=\text{æm} > \text{ketab}.[\text{ha.jæm}]$ book-PL=1.SG (glide insertion)

b. $\text{ba}=\text{æm} > [\text{ba.hæm}]$ with-1.SG (h-insertion)

c. $\text{mi-xa}=\text{æm} > [\text{mi.xam}]$ DUR-want=1SG (vowel deletion)

(5) Kinnauri (Sino-Tibetan; www.linfil.uu.se/personal/anjusaxena/kintext07.html)

a. $\text{byo-o} > [\text{bi.jo}]$ or $[\text{bo:}]$ go-PROG (glide-insertion or coalescence)

b. $\text{bə-o} > [\text{bə.do}]$ come-PROG (consonant insertion)

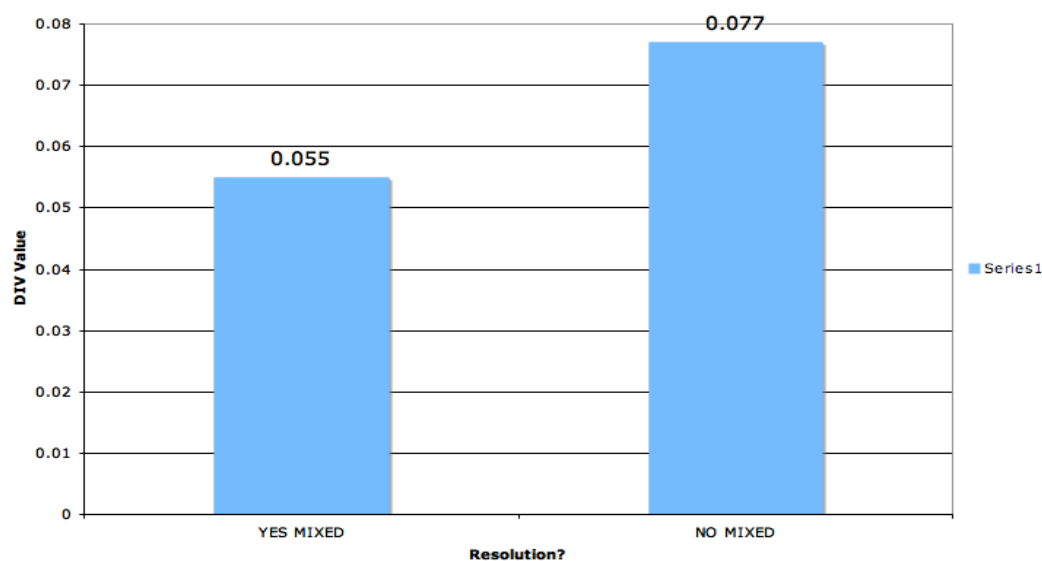
Question 6: Given that there are several languages with mixed strategies in a single domain, would we see a similar pattern/degree of non-isomorphism (Diversity) of P-words in the same language?

- DIV (Bickel & Hildebrandt 2005): The number of non-isomorphic P-words in a language; a value approaching .1 indicates higher degree of non-isomorphism; a value approaching .00 indicates a lower degree

(6) Chukchi (Chukotko-Kamchutkan) (Partial Representation) (DIV = .095)

	CF1	PF	Σ	CF2	SF	ENCL
Vowel-Harmony						
Vowel Glottalization						
Nasal Coda P.O.A.						

Fig. 6 Average DIV Values



Answer 6: Languages with mixed resolution types within a single domain have a generally LOWER DIV value, while languages with non-mixed types have a HIGHER DIV value; however, these findings are not significant.

- Although we cannot say with certainty, the suggestion is that mixed resolution strategies are not necessarily indicative of more non-isomorphism.

ii. ± Resolution, Same Domain

- A language may or may not resolve hiatus cases within the same domain type.
- Dolakha Newar: (C)(Cg)V(V)(C) (Genetti 1994); A dedicated section on “vowels in combination”

(7) Stem-Suffix in Dolakha Newar

a. Vowels of identical quality

bi-i > [bi] give-INF (vowel elision)

b. Stem final high vowel & suffix high vowel

bi-u > [bju] give-NR1 (glide formation)

c. Exception: nominalizer NR2 -e; the result of this hiatus is “a disyllabic word of two distinct pulses” (p. 30)

ye-e > [je.e] come-NR2 (heterosyllabification)²

d. Exception: -eu 3sFUT or -en PART following stem-final /e/, also two distinct pulses, but vowel quality change

ye-eu > [ye.ɛu] come-3s.FUT (heterosyllabification)

- Cases like these not dealt with in OT or Autosegmental or domains literature
- We must turn to both phonology & item-based information to account for ± resolution & type

² I don't have data, but I assume that other -e suffixes like the locative and the linker are subject to deletion when following stems with identical vowel quality

IV. Concluding Remarks

- Resolution of hiatus is cross-linguistically preferred, seems to characterize stocks
- A focus on resolution types within theoretical framework overlooks language-specific particularities & complexities
- And at this point it is not likely that the facts about a particular process in a language are necessarily indicative of the overall prosodic (word) profile of that language
- Additional support for individual historical sources of observed phenomena

References

- Acharya, J. 1991 *A descriptive Grammar of Nepali and an analyzed corpus*. Georgetown University Press. Washington, D.C.
- Bickel, B., and K. A. Hildebrandt. 2005. "Diversity in Phonological Word Domains: Testing the Prosodic Hierarchy Hypothesis". 6th International Conference of the Association for Linguistic Typology.
- Blevins, J. 1995. The Syllable in Phonological Theory, *Handbook of phonological theory*, ed. by John Goldsmith. London: Blackwell, 206-44.
- Casali, R. 1997. Vowel elision in hiatus contexts: Which vowel goes? *Language* 73.3 493-533.
- Casali, R. 1995. *Resolving hiatus*. UCLA Dissertation. ROA #215-0997.
- Donwa-Ifode. S. 1985. Glide formation, elision, assimilation and contraction: A reassessment Evidence from Isoko. *Journal of West African languages* 15.41-55.
- Genetti, C. 1994. *A Descriptive and Historical Account of the Dolakha Newari Dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa (*Monumenta Serindica*, 24).
- Hyman L. M. & Francis X. Katamba. *The Word in Luganda*, Ms.
- Langdon, M. 1970. *A grammar of Diegueño: The Mesa Grande Dialect*. University of California Press.
- Maddieson, I. 2005. Syllable Structure. *The world atlas of language structures*. eds. Martin Haspelmath, Matthew S. Dryer, David Gil, Bernard Comrie. Oxford, UK: Oxford University Press.
- Mahootian, S. 1997. *Persian*. London: Routledge.
- Schane, S. The resolution of hiatus. *CLS 23: Papers from the 23rd annual regional meeting of the Chicago Linguistic Society. Part 2: Parasession autosegmental and metrical phonology*. 279-290.
- Watters, D. E., with the participation of Yogendra P. Yadava, Madhav P. Pokharel, Balaram Prasain. 2005. *Notes on Kusunda Grammar: a linguistic isolate of Nepal*, Kathmandu: National Foundation for the Development of Indigenous Nationalities
- Watters, D. E. 2002. *A Grammar of Kham*. Cambridge University Press.

APPENDIX: LANGUAGE SAMPLE

LANGUAGE	HIATUS POSSIBLE?	HIATUS RESOLVED?	GEN. AFFIL.	AREA	SPHERE (if S-T)
Belhare			Sino-Tibetan	Indic	Indosphere
Boumaa Fijian			Austronesian	Oceania	
Cantonese			Sino-Tibetan	SE Asia	Sinosphere
Cayuvava			Isolate	NE S America	
Chukchi			Chukotko-Kamchutkan	N Coast Asia	
Diegueño			Yuman	California	
Diyari			Pama-Nyungan	S. Australia	
Dolakha Newar			Sino-Tibetan	Indic	Indosphere
Egyptian Arabic			Semitic	N. Africa	
Finnish			Uralic	Europe	
German			Indo-European	Europe	
Greek (Modern)			Indo-European	Europe	
Hixkaryana			Cariban	NE S. America	
Irish			Indo-European	Europe	
Jahai			Austro-Asiatic	SE Asia	
Kayah Li			Sino-Tibetan	SE Asia	Sinosphere
Kewa			Engand	S. New Guinea	
Kinnauri			Sino-Tibetan	Indosphere	
Kalkha Mong.			Mongolian	Inner Asia	
Kham			Sino-Tibetan	Indic	Indosphere
Kusunda			Isolate	Indic	
Kyirong Tibetan			Sino-Tibetan	Indic	Buffer Zone
Lahu			Sino-Tibetan	SE Asia	Sinosphere
Limbu			Sino-Tibetan	Indic	Indosphere
Lithuanian			Indo-European	Europe	
Luganda			Benue-Congo	S. Africa	
Manange			Sino-Tibetan	Indic	Buffer Zone
Mandarin			Sino-Tibetan	SE Asia	Sinosphere
Mapudungun			Isolate	Andean	
Martuthunira			Pama Nyungan	S. Australia	
Meithei			Sino-Tibetan		Buffer Zone
Mundari			Austro-Asiatic	Indic	
Nama			Kwadi-Khoe	S. Africa	
Nepali			Indo-European	Indic	
Persian			Indo-European	Cauc.-Mesop.	
Polish			Indo-European	Europe	
Sko			Macro-Sko	W. N. Guinea	
Santali			Austro-Asiatic	Indic	
Spanish			Indo-European	Europe	
Swedish			Indo-European	Europe	
Turkish			Turkic	Europe	
Vietnamese			Austro-Asiatic	SE Asia	
Wu (Chinese)			Sino-Tibetan	SE Asia	Sinosphere
Yidiny			Pama-Nyungan	N. Australia	
Yimas			L. Sepik	Inn N. Guinea	