

Bergische Universität Wuppertal:
Zentrum für interdisziplinäre Sprachforschung

**Complexity between
Explicitness and Economy—
Overt vs. Hidden Complexity**

Walter Bisang (Mainz)

Wuppertal, 19. Dezember 2012



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Survey

- 1. Two types of complexity: Overt complexity vs. hidden complexity (section 1)**
- 2. Hidden complexity is significantly higher in Chinese (and other East and mainland SE Asian languages, EMSEA) than in other languages: Comparison with West African and Creole languages (sections 2 and 3)**
- 3. This high degree of hidden complexity reflects an economy-oriented process of maturation (cf. Dahl 2004 and his explicitness-based approach) (section 4)**

1.

**Two Types of Complexity:
Overt Complexity vs.
Hidden Complexity**

Complexity as it is currently discussed I

I. Hidden simplicity:

Recursion:

- **Faculty of Language—Narrow Sense:**

Basic property of human language that is uniquely human (Hauser, Chomsky & Fitch 2002)

- **Typological approach (Givón & Shibatani 2009)**

II. Overt complexity:

Surface-oriented:

How are grammatical categories overtly marked?

Complexity as it is currently discussed II

Complexity in terms of McWhorter (2001, 2005)

[A]n area of grammar is more complex than the same area in another grammar to the extent that it encompasses more overt distinctions and/or rules than another grammar. (McWhorter 2005: 45)

Four diagnostics, diagnostic (iii):

A grammar is more complex than another to the extent that it gives overt and grammaticalized expression to more fine-grained semantic and/or pragmatic distinctions than another. (McWhorter 2005: 45)

There is a third type of complexity:
Hidden Complexity

What needs to be expressed in language I (Sapir 1921)

(1) *The farmer kills the duckling.*

**Categories that need to be expressed in English
(Sapir 1921: 82):**

- Number**
- Definiteness**
- Tense**
- Person/number agreement of subject
on the verb**

What needs to be expressed in language II (Sapir 1921)

In Chinese, none of these categories needs to be expressed (Sapir 1921: 90, the brackets are mine):

(2) (*Farmer*) kill (*duck*).

Languages differ essentially in what they must convey and not in what they may convey.

Jakobson (1992 [1959]: 49)

Reasons for differences in terms of obligatory categories I

Necessity of pragmatic enrichment:

Morphosyntactic structures never fully express the meaning they have in a concrete situation.

Reason:

Human speech is characterized by its extremely **slow transmission rate**.

Prearticulation, parsing and pragmatic inference can assess and generate much more information in much less time (Levinson 2000).

Reasons for differences in terms of obligatory categories II

The articulatory bottleneck:

**“inference is cheap, articulation expensive”
(Levinson 2000: 29)**

Reasons for differences in terms of obligatory categories III

There are two types of forces that take influence on the grammar of languages:

Articulation

Pragmatic Inference



Explicitness



Economy

**Competing motivations (Gabelentz 1891,
Haiman 1983, Optimality Theory).**

Competing motivations — two types of complexity

Explicitness: Overt complexity

The structure of the language forces the speaker to explicitly encode certain grammatical categories.

The relevant categories are morphosyntactically **overt** and can be measured.

Economy: Hidden complexity (Bisang 2009)

The grammar of a language does not force the speaker to use a certain grammatical category.

The grammatical category is **hidden**. Access only through inference.

Hidden complexity I

Hidden complexity shows up in two different forms:

Omission of grammatical markers (cf. §§ 2, 3):

- Zero arguments (radical pro-drop)
- Multiple coreference in relative clauses
- Clause combining: Lack of overt markers

Multifunctionality:

(will not be discussed in detail here, cf. Bisang 2009 for some examples).

2. Hidden Complexity in Chinese

Case 1:

Zero arguments - radical pro-drop

Rizzi's (1986) classical account for pro-drop:

- Formal licensing (by a head)
- Licensing by content through rich agreement

Problem: Radical pro-drop:

(4) Chinese:

你看了电影吗? — 看了。

Ni kan-le dianying ma? \emptyset kan-le \emptyset .

you see-PF film Q saw-PF

‘Did you see a film?’

‘[I] saw [one].’

Case 1:

An example from a real text:

(5) *Zixia de tudi Gongsun Gao lai zhao*
Zixia GEN disciple Gongsun Gao come look

Mozi. yijing haoji huile, zongshi
Master Mo already.several.times never

ø bu zai jia. ø jian bu zhao ø.
NEG be.at home meet NEG can

‘**Gongsun Gao**, a disciple of Zixia, was looking for **master Mo** several times and [**he**] was never at home, so [**he**] was unable to meet [**him**].

(LuXun 鲁迅, Feigong 非攻 ‘Opposing Aggression’.)

Case 2:

Relative clauses and coreference I

The syntactic function of the head noun in the relative clause must often be inferred from context.

(14) [找的]人還沒有回來。

[zhao de] ren hai meiyou hui-lai.

look.for REL man still NEG:PF return-come

a. Object coreference:

‘The people [(we) were looking for]
haven’t come back yet.’

b. Subject coreference:

‘The people [who looked (for us)]
haven’t come back yet.’

Case 2: Relative clauses and coreference II

Unmarked coreference with non-arguments:

(15a) **Possessor:**

na ge [tofa hen zhang de] xuesheng
that CL hair very long REL student
‘the student whose hairs are very long’

(15c) **Instrumental:**

[wo xie xin de] maobi
I write letter REL pencil
‘the pencil [with which I write letters]’

Case 3:

Clause combining

Omission of adverbial subordinators. The construction represented by a given surface structure must be inferred from context.

(23) Li & Thompson (1981: 595):

Wo [mai piao] [jin-qu].

I buy ticket enter-go

a. **Purpose:**

‘I bought a ticket to go in.’

b. **Consecutive action:**

‘I bought a ticket and went in.’

3.

**Hidden Complexity in
West Africa and Creoles**

West African: Pro-drop in Yoruba I

In Yoruba main clauses, there must always be an overt subject and an overt object (also cf. Fongbe in Handout, p. 5):

(A) a. Ayo rà aso.
Ayo HTS buy cloth
'Ayo bought clothes.'

b. Ó rà aso.
3.SG buy cloth
'He bought clothes.'

c. Ó rà á.
3.SG buy 3.SG:OBJ
'He bought it.'

West African: Pro-drop in Yoruba II

If one or both positions are zero, the sentence is ungrammatical:

d. * \emptyset rà aso.
buy cloth

e. * Ayo ' rà \emptyset .
Ayo HTS buy

f. * \emptyset rà \emptyset .
buy

Yoruba is definitely not a radical pro-drop language.

Relative Clause Formation: Yoruba I

(also cf. Fongbe, Handout p. 7, ex. (16))

N [ti ... V ...]

Subject coreference: obligatory subject pronoun:

(B) *obìnrin* [t'ó maa ràá]

woman REL'3.SG:SUBJ TAM buy:OBJ.3.SG

‘the woman [who bought it]’

Relative Clause Formation: Yoruba II

Object coreference: the object pronoun is zero
(a form which is ungrammatical as a matrix clause):

(C) *iṣu_i [tí mo rà ø_i lánàá]*
yam REL 1.SG buy yesterday
'the yam [I bought yesterday]'

Locative coreference: with locative particle *tí*:

(D) *ó mò ọjà [tí mo tí ràá]*
3.SG know market REL 1.SG LOC buy:3.SG.OBJ
'He knows the market [where I bought it].'



Different obligatory structures for different types of coreference. No options for inference.

Clause combining in Yoruba I

Different semantic relations between clauses require different constructions:

Purpose:

(25a) *Mo ra tíkéètì láti wole.*
1.SG buy ticket PURP enter
‘I bought a ticket to go in.’

Consecutive action:

(25b) *Mo ra tíkéètì, mo sì wole.*
1.SG buy ticket 1.SG and enter
‘I bought a ticket and went in.’

Clause combining in Yoruba II

Mere juxtaposition of two events does not trigger pragmatic inference—it iconically reflects conceptual closeness. Conceptual closeness can be established in discourse.

[How did you get access to the soccer game? —
By knowing the governor or by buying a ticket?]

(24) *Mo ra tíkéètì wole ni.*

1.SG buy ticket enter FOC

‘I entered by buying a ticket.’



The grammar systematically blocks surface structures which may refer to more than one construction of clause combining.

Summary:

Chinese and West African (Yoruba)

To what extent does the grammar of a language leave options for pragmatic inference of grammatical information?

	Chinese	Yoruba
Pro-drop	a lot	no options for inference
Rel Clause	a lot: no subject/object asymmetry	no options for inference
Clause Comb	non-marking provides options for inference	non-marking has a specific meaning, no options for multiple analysis

General remarks on Creoles

Creole languages I have looked at:

Atlantic Creoles:

- **Angolar:** Portuguese-based, spoken on São Tomé, an island in the Gulf of Guinea (Africa)
- **Berbice Dutch:** Dutch-based, spoken in Guyana, S America
- **Haitian Creol / Kreyòl Ayisiyen,** French-based, spoken in Haiti

Non-Atlantic Creoles:

- **Mauritian Creole:** French-based, spoken on Mauritius Island (SW Indian Ocean),
- **Tok Pisin:** English-based, a variety of Melanesian Pidgin, spoken in Papua New Guinea
- **Zamboangueño:** Spanish-based, a variety of Chabacano (Philippines) on the Southern tip of Mindanao Island

Overt complexity and Creoles

Creoles represent the world's most simple languages (McWhorter 2001, 2005).

They did not have enough time to develop more complex overt structures as we find them in older languages.

Creoles, in being recently borne of communication vehicles deliberately designed to eschew all but the functionally central (pidgins), are unique examples of natural languages with much less contingent accumulation of “ornamental” elaboration than older grammars drag along with them. (McWhorter 2005: 43)

Creoles I

Pro-drop:

is not as developed in most creoles as in Chinese.

Exceptions: Languages in which the contact languages involved allow pro-drop.

Relative clauses:

There is at least a distinction between coreference with arguments vs. non-arguments.

Exceptions are again motivated by the properties of the contact languages involved.

Clause combining:

The use of adverbial subordinators is much more like in Western languages and in Yoruba than in Chinese.

Creoles II: Pro-drop in Haitian

Haitian Creole does not seem to be a pro drop language (Déprez 1994, Neeleman & Szendrői 2007):

- (15) a. ***Li*** *pati.* b. **pati.*
 3.SG leave leave

‘S/He left.’ (DeGraff 1993: 72)

- (16) **(Li) difisil pou pale ak Jan.*
 3.SG difficult to talk with John

‘It is difficult to talk to John.’ (Déprez 1994)

- (17) \emptyset *genle* *Jak* *damou*.
 seem Jack be.in.love

‘It seems that Jack is in love.’ (DeGraff 1993: 72)

Creoles III: Pro-drop in Zamboangueño

Radical pro-drop:

Substrate: Philippine languages with radical pro-drop

Superstrate: Spanish with (mild) pro-drop

Lipski & Santoro (2007: 376):

(12) *Andá ϕ alyì na réyno*
go there to kingdom
‘[He] goes there to the kingdom’

(13) *Kwándo sale ϕ afwéra ya murí ϕ .*
when leave outside PST die
‘When [he] went outside, [he] died.’

Creoles IV: Relative clauses in Angolar

Different relative markers for subject vs. non-subject coreference:

Subject coreference (Maurer 1995: 55):

- (17) *ome si [ki ba tamba]*
man DEF REL go catch.fish
'the man [who left to catch fish]'

Creoles V: Relative clauses in Angolar

Object coreference (Maurer 1995: 55):

- (18) *ome si [ma m bê]*
man DEF REL 1.SG see
'the man [I saw]'

Non-argument coreference (Maurer 1995: 55):

- (19) *ome si [ma n ga taba ku ê]*
man DEF REL 1.SG TAM work with 3.SG
'the man [with whom I work]'

Creoles VI: Relative clauses in Zamboangueño

Subject coreference (Lipski & Santoro 2007: 383):

- (20) *el mana hénte [kyén ya man tunúk ø*
DEF PL people REL PST DRV be.prick
na gargánta]
in throat
‘people who have gotten fish spines caught
in their throats’

Creoles VII: Relative clauses in Zamboangueño

Non-argument coreference (Lipski & Santoro 2007: 383):

(22) *El persona [kon-kyen ta kombersá tu]*
DEF person with-REL PROG talk you

byen bwéno gayót.
very good EMPH

‘the person [you are talking to] is very nice indeed.’

Creoles VIII: Clause combining

Finding good data is not easy. More like European languages. Simple dropping of the conjunction leads to different interpretation:

Berbice Dutch: Conditional (Kouwenberg 12994: 115):

(29) *Aso wa krik-it-o hiri, o wa*
COND PST get-PF-3.SG here 3.SG PST
bato doto.
kill:PF.3.SG dead

PST + PF with *aso* 'if' => hypothetical meaning

(29a) 'If he had gotten it here, it would have killed him.'

PST + PF without *aso* => Anterior meaning

(29b) '[?] He had gotten it here. It killed him.'

Conclusion I

To what extent does the grammar of a language leave options for pragmatic inference of grammatical information?

	Chinese	Yoruba	Creoles
Pro drop	a lot	no options for inference	some options for inference
Rel Clause	inference of subj/obj co-reference	no options for inference	inference mostly not necessary
Clause Comb	no marking provides options for inference	non-marking has a specific meaning	non-marking provides limited options for inference

Conclusion II

Creoles tend to reduce grammatical categories to those which are absolutely necessary for communication (cf. McWhorter 2001, 2005).

In spite of this,
their grammatical systems are less open to the
pragmatic inference/hidden complexity than
those of Chinese (and EMSEA languages).



The grammatical systems of Creoles represent the limits of what can be omitted from the perspective of the grammars of the substrate (West African) and superstrate (European) languages involved.

4.

**Economy-Based
Maturation, the Case of
EMSEA Languages**

Explanations I:

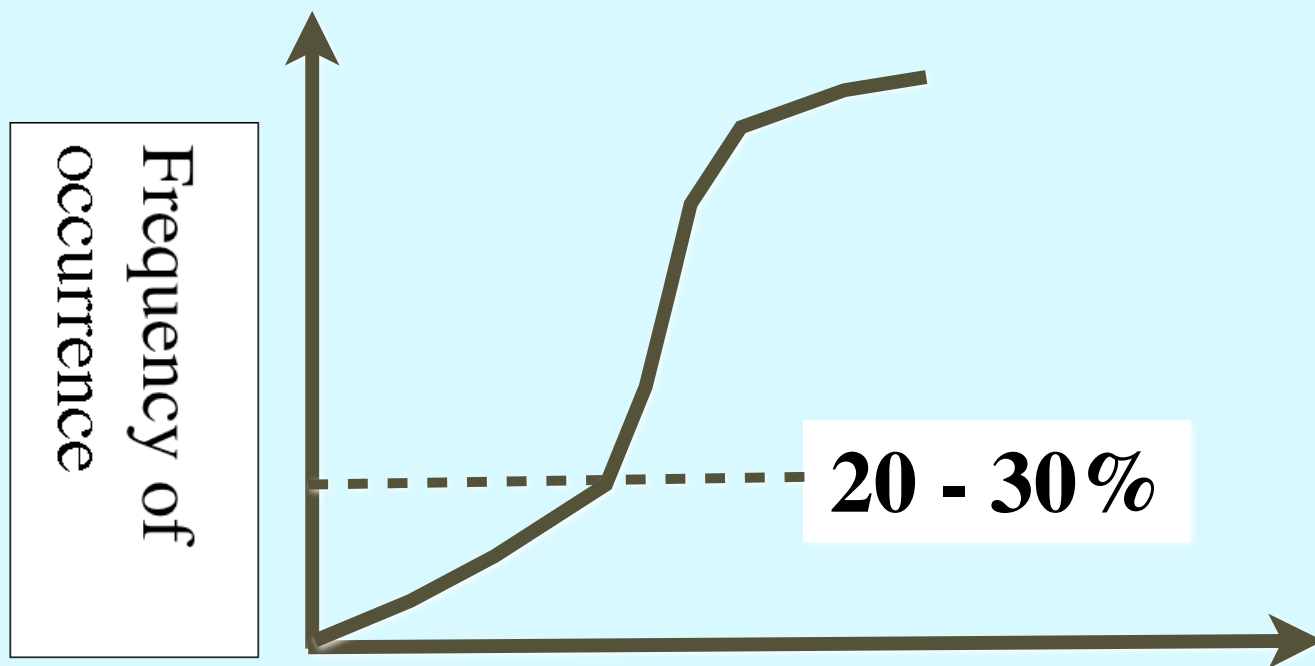
Basic diachronic question:

What favored the high degree of hidden complexity in EMSEA languages?

Frequency and its role in language change and the factors that keep it low:

- (i) Input (what is already there)**
- (ii) No morphological expression of inflectional category**
- (iii) Language contact as a sociocultural factor**

Explanations II: Frequency and successful language change



Time axis: degree of diffusion within a speech community

For a change to be successful, it needs to reach 20 - 30% frequency in the relevant population of utterances (Wang & Cheng 1970, Bailey (1973)).

Explanations III:

Frequency and successful language change

The S-curve model with its 20 - 30% frequency also applies to changes that lead to:

- Radical pro-drop —> Non-pro-drop
- No overt coreference distinctions in RC —> obligatory coreference distinctions
- —> obligatory marking in clause combining



To understand what prevents economy from developing beyond a certain degree, it is necessary to understand what factors keep frequency low.

**Explanations IV:
What is already there and inflectional
morphology in Chinese**

**See
Handout, pp. 12 - 13**

**Explanations V:
What is already there and inflectional
morphology in Niger-Congo**

**See
Handout, pp. 13 - 14**

Explanations VI: Language contact

EMSEA languages are generally characterized by their high degree of hidden complexity.

At the time when they got in contact, the non-Sinitic languages had a similarly low degree of morphology:

- Tai-Kadai (e.g. Thai)
- Hmong-Mien (e.g. Hmong)
- Mon-Khmer (Vietnamese, Khmer/Cambodian)
- [Chamic (Austronesian)]

Explanations VII: Language contact

In some Mon-Khmer languages, however, morphology had the potential to become productive.

=> Example: Nominalization in Khmer

(33) The infix **-m-** in Khmer: Agent nouns:

sò:m ‘ask’	—> s-m-ò:m ‘beggar’
cam ‘wait for, guard’	—> ch-m-am ‘guardian’
cù:əp ‘do business’	—> ch-m-ù:əp ‘business-man’

Explanations VIII: Language contact

But other contact languages have developed alternative, syntax-derived morphological devices. These devices became productive in Khmer, too.

(34) The head noun **nèək** ‘man’ in Khmer:

nèək-daə [person-walk] ‘pedestrian’

nèək-taɛŋ [person-compose/write]

‘author, composer, writer’

nèək-chlò:p [person-go stealthily to watch someone] ‘spy’

Maturation and overt complexity I

McWhorter (2001, 2005) on the accumulation of overt complexity as a result of long-term history.

Creoles, in being recently borne of communication vehicles deliberately designed to eschew all but the functionally central (pidgins), are unique examples of natural languages with much less contingent accumulation of “ornamental” elaboration than older grammars drag along with them. (McWhorter 2005: 43)

Maturation and overt complexity II

Similarly, Dahl (2004) on maturation:

Maturation:

The accumulation of material in a grammar G that did not exist at an earlier stage G' of the same language.

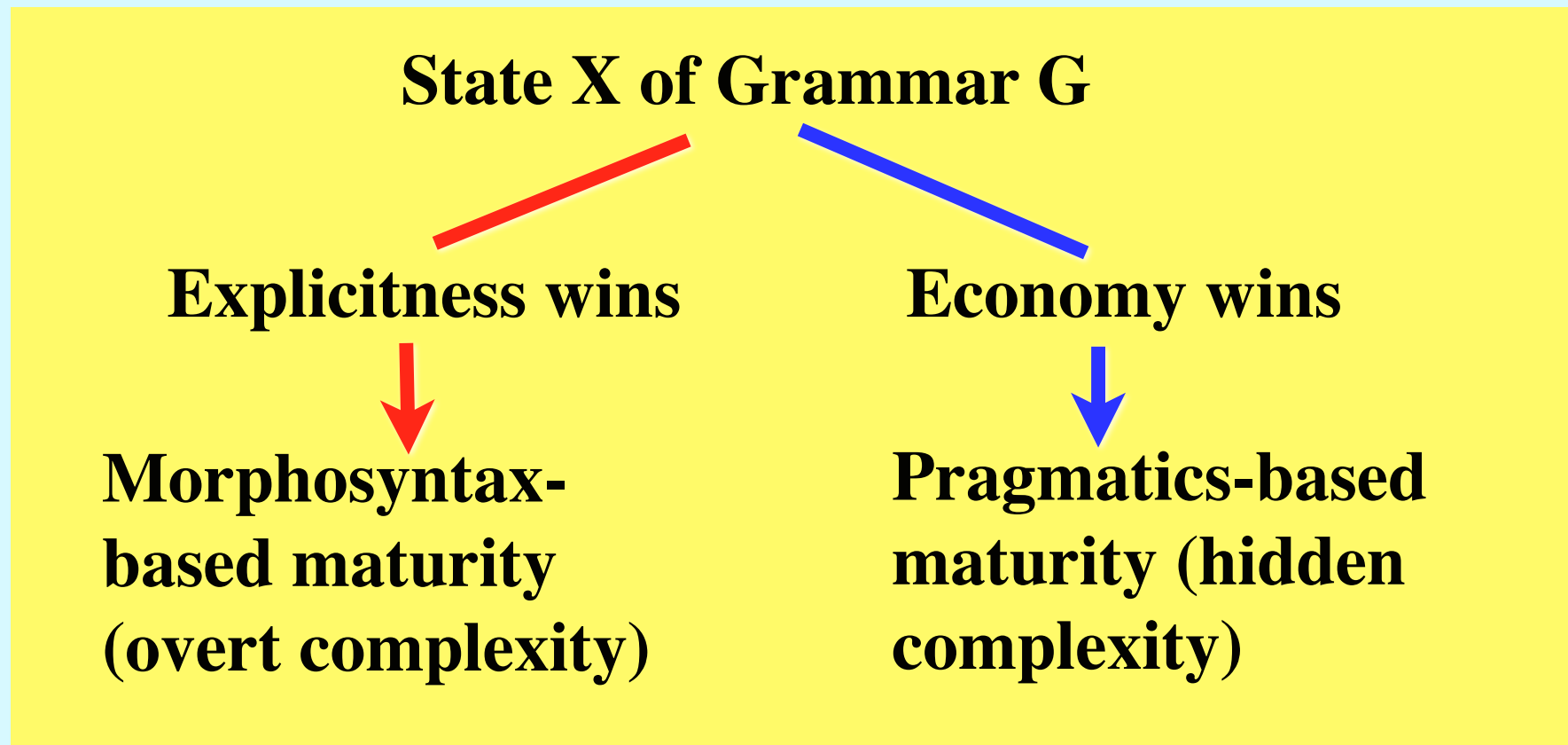
Maturation and overt complexity III

This type of maturation generates phenomena such as the following (Dahl 2004: 114-115):

- **Complex word structure** (inflectional, derivational and incorporating morphology)
- **Lexical idiosyncrasies** (grammatical gender, inflectional classes, idiosyncratic case marking)
- **Syntactic phenomena that are dependent on inflectional morphology** (agreement)
- **Specific marking of subordinate clauses**
- **Morpheme and word level features in phonology**

Maturation and hidden complexity IV

Maturation cannot only follow the direction of explicitness, it can also follow economy:



Maturation and hidden complexity V

In Chinese (and other EMSEA languages), economy wins to the extent that it has reached a degree of pragmatics-based maturity (or hidden complexity) which is rarely seen elsewhere.

Omission:

- Radical pro-drop
- Multiple coreference options in relative clauses
- Optionality of adverbial subordinators

Multifunctional grammatical markers:

- (in)definite classifiers in [CL-N]

5. Conclusions

In this paper, I tried to show that

- There are two sides of complexity: overt complexity (explicitness) and hidden complexity (economy).
- High degree of hidden complexity in EMSEA languages: radical pro-drop, coreference in RCs, clause combining.
- In most Creoles, hidden complexity is lower than in EMSEA languages. Creoles represent the limit of what can be reduced/omitted, given the input languages.
- High hidden complexity in EMSEA is due to the joint impact of three frequency blocking factors.
- Economy-oriented maturation is a second type of maturation that is prominent in EMSEA languages.

謝謝你
Thank you!

Zentrum für interdisziplinäre
Sprachforschung
Wuppertal, 19.12.2012



JOHANNES GUTENBERG
UNIVERSITÄT MAINZ

Structure of the remaining paper

2. Hidden complexity in Chinese:

- (i) radical pro-drop, (ii) coreference in relative clauses, (iii) clause combining

3. Hidden complexity: Comparison with:

- Creole languages (low overt complexity)
- West African languages (important substrate languages for many Atlantic Creoles).

4. Explanation: economy-oriented maturation

Factors that enhance hidden complexity in EMSEA:

- (i) Input (what is already there), (ii) morphology that does not express inflectional categories, (iii) language contact.