Pidgin Genesis in Optimality Theory[1]

Joan Bresnan

Stanford University

Proceedings of the LFG98 Conference
The University of Queensland, Brisbane

Miriam Butt and Tracy Holloway King (Editors)

1998
CSLI Publications
http://www-csli.stanford.edu/publications/

[1]This revised version of the paper given at Lfg98 contains significant changes and corrections; it supersedes the version appearing in the printed LFG98 Proceedings. Many thanks to Sarah Julianne Roberts and Elizabeth Traugott for helpful comments and suggestions on earlier drafts. I am also grateful to Bill Foley, John Rickford, Aaron Broadwell, and to Peter Bakker, Salikoko Mufwene, Claire Lefebvre and other members of the CreoLIST (http://www.ling.su.se/Creole/) who generously answered my questions about pidgins and freely offered their materials and expertise. None of them should be blamed for my mistakes.

Pronouns in pidgins

Pidgins arise in certain contact situations as a conventionalized means of communication between groups of adult speakers of different vernacular languages (Thomason 1997). Pidgins have a diversity of structural forms often including some unusual or marked structures from their source languages, as well as recurring universal properties, such as the preponderance of analytic syntax, CV syllable structure, and generic lexical semantics (Thomason and Kaufmann 1988, Foley 1988, Holm 1989, Bakker 1994). In particular, there is a vast prevalence of free pronouns in pidgins. Why should this be?

(1) "Pidgins prefer free pronoun forms to bound ones." —Mühlhäusler and Harré (1990: 262)

One hypothesis is that pidgins simply derive their pronominal forms from their source languages. A version of this hypothesis is stated in (1):

(2) Hypothesis I: Pidgin pronouns resemble those of the language that provides most of their lexicon (their lexifier).

This is a very natural hypothesis because so many of the languages which provide the lexicons of well-known pidgins are European, arising from European exploration, missionary settlement, trade, colonization, plantation agriculture, commercial whaling expeditions, and the like. Thus it is not surprising that the English-lexifier pidgin of New Guinea, Tok Pisin, employs freestanding pronouns, because English employs free pronouns. The same holds for the Ndyuka-Trio Pidgin, a contact language of Suriname used by the Ndyuka (a "Bushnegro" society) and the Trio Indians. The syntax of Ndyuka-Trio Pidgin closely follows that of the indigenous Indian language, while the larger part of its lexicon, including its freestanding pronouns, comes from the Ndyuka's language, which is an English-lexifier creole (Huttar and Velantie 1996). These better and lesser-known instances of pidgins (among others) support Hypothesis I.

However, as Haiman (1985: 161–2) notes, Hypothesis I does not explain why West African Pidgin Portuguese (according to Naro (1973: 444)) used the fully stressed, independent strong pronominal forms where the various clitic pronominals were used in Portuguese. Haiman (1985: 161ff) also cites Bantu pidgins/creoles such as Kenya Swahili and Fanagalo as evidence against Hypothesis I, to which could be added Kituba (Mufwene 1996). However, uncreolized pidgins will be of primary interest here, because (as we will see below) it is the absence of native speakers that poses the conceptual problem to be addressed. Further, it appears that creoles develop richer pronominal types than pidgins, including clitic or bound pronominals (DeGraff 1992, Deprez 1992).

The inadequacy of Hypothesis I becomes clear when we examine various pidgins whose lexicons come from (non-European) languages with bound pronominal systems. See Table 1.[2]

[2]This table is based on the following sources: Foley 1988, n.d., and Williams 1993 on Yimas Pidgin; Nichols 1995 on Broken Oghibbeway; Drechsel 1997 on Mobilian Jargon; Dutton 1996 on Hiri Motu; van der Voort 1996 on Eskimo Pidgin; Thomason and Kaufmann 1988, Samarin 1996, and Thomason 1997 on Chinook Jargon; Goddard 1996 on Pidgin Delaware; Roberts 1995 and p.c. May 1998 on Pidgin Hawaiian.

(3) Table 1: Pidgins with bound-pronominal lexifier languages:

Pidgin	Lexifier	Other source languages
Yimas Pidgin	Yimas (Papuan)	Arafundi, Alamblak, other Papuan
Broken Oghibbeway (early 19th c.)	Ojibwe (Algonquian)	Wisconsin Amerindian, and European?
Mobilian Jargon (late 17th to 20th c.)	Choctaw, Chickasaw (Muskogean)	S.E. Amerindian
Hiri Motu	Motu (Austronesian)	Papuan, Austronesian, and European
Eskimo Pidgin	W. Greenlandic	European
Chinook Jargon (19th and 20th c.)	Chinook, Nootka (Chinookan, Wakashan)	N.W. Coast Amerindian and French
Pidgin Delaware (17th c.)	Unami Delaware (E. Algonquian)	Dutch
Pidgin Hawaiian (late 18th to early 20th c.)	Hawaiian	European, Cantonese

All of these pidgins have free pronouns.

(4) shows an example from the pidgin called "Mobilian Jargon" compared to its lexifier Choctaw, spoken in the Southeastern United States (Drechsel 1997: 300):

(4) Mobilian Jargon: 'I want water./I am thirsty.'
oka eno banna
water I want

Choctaw: 'I am thirsty.'
oka sa-banna-h
water 1SG-want-PREDICATIVE

Mobilian Jargon uses freestanding syntactic pronouns where Choctaw uses bound pronominals.

(5) compares Pidgin Hawaiian with its Hawaiian lexifier. The possessive pronominals of Hawaiian can occur either postnominally as analytic pronouns or prenominally, bound to the definite article. These pronouns express alienable/inalienable distinctions through the thematic vowel *o/a*. In Pidgin Hawaiian, however, only the freestanding pronoun occurs in possessives, and it lacks case or alienable/inalienable distinctions.

(5) Pidgin Hawaiian: 'your hat'
kela papale oe
DEF hat you

Hawaiian:
ka pa:pale a:u
DEF hat ALIENABLE.you.POSS
k-a:u pa:pale
DEF-ALIENABLE.you.POSS hat

Further, languages with bound pronominals have the typological property that their free pronouns appear to be specialized for focus uses (Schwartz 1986, Bresnan and Mchombo 1987, Givón 1983). But pidgins based on such languages employ the free pronouns of the lexifier in the contexts where bound pronouns would be used.

Consider (6) comparing Yimas Pidgin with its lexifier Yimas, a Papuan language of New Guinea (Foley 1988: 171):

(6) Yimas Pidgin: 'I hit him'
Ama min namban kratiki-nan.
1SG 3SG toward hit-NONFUT

Yimas:
Na-ka-tupul.
3SGO-1SGS-hit

In Yimas Pidgin free subject and object pronouns are used where bound pronominals are used in Yimas. Note that the Yimas Pidgin pronouns are cognate with Yimas *ama* 1SG and *m-n* (3 DEICTIC). These free forms are used used 'contrastively' in Yimas, according to (Foley 1991: 112).

Similarly, in Pidgin Delaware, according to Goddard (1996: 57),"[t]he three most generally used pronouns reflect the first and second singular of the Unami [Delaware–jb] emphatic pronoun set and an emphatic form of the inanimate singular proximal deitic".

While Hypothesis I is can thus be rejected for not explaining the difference in uses of free pronouns in pidings and their lexifiers, a closely related hypothesis suggested by Patrick McConvell in a posting to the CreoLIST (McConvell 1997a) fares much better:

(7) Hypothesis II: Free pronouns occur in pidgins through contact with languages lacking bound pronouns.
"In all cases that I know of, the language/dialect which has lost bound pronouns abuts onto languages without bound pronouns. The loss is related to some contact phenomenon, pretty certainly — the question is what exactly?" —Patrick McConvell, CreoLIST, May 28, 1997

This hypothesis fares better because most pidgins in Table 1 have contact languages lacking bound pronominal systems. However, Hypothesis II still would not account for Yimas Pidgin, where the contact languages all have bound-pronominal systems (as McConvell 1997a,b also observes); see Foley 1988, n.d. Furthermore, although documentation is scanty, Mobilian Jargon has also been argued in recent research to pre-date European influences, being used as a contact language between Amerindian groups (Drechsel 1997: 274–286).

Hypothesis II also faces the conceptual problem of explaining exactly what the mechanism is that produces free pronouns by contact (as McConvell notes in the quotation in (7)). The solution is not at all obvious, especially since the contact language lacking bound pronouns need not provide either the free pronouns or the morphosyntax of the pidgin, as the case of Delaware Pidgin shows. Goddard 1996 gives evidence that Pidgin Delaware arose from contact between Eastern Algonquian Delaware and Dutch, but has free pronouns reflecting the emphatic pronouns of Unami Delaware, as noted above (Goddard 1996: 67).

A third approach, suggested by a number of authors over the years (e.g. Kay and Sankoff 1974, Heine 1975, Mühlhäusler 1986, Haiman 1985), is the universalist hypothesis, represented in both formalist and functionalist versions in (8):

(8) Hypothesis III (formalist version): Free pronouns represent the default parameter setting of Universal Grammar, which characterizes the initial state of the language learner. Stable pidgins have free pronouns because they reflect the initial state of the language learner.

Hypothesis III (functionalist version): Free pronouns are unmarked pronominal forms crosslinguistically. The isolating, analytic, uniform syntactic structures of pidgins can be explained in terms of their extreme syntactic unmarkedness, which facilitates learning.

The problem for Hypothesis III is the existence of bound pronominals in some pidgins. (9) compares the nineteenth-century pidgin called "Broken Oghibbeway" (Nichols 1995) with its lexifier Ojibwe, an Algonquian language of North America. The examples are from Bakker (1995: 31–2):

(9)

Broken Oghibbeway: 'He fears me.'		Ojibwe:
O-kot-aan	*niin.*	*Ni-gos-ig*
3SG.AN-fear-3.INAN	1SG	1SG-fear-INV.3SG.SUBJ

While Broken Ojibwe has a freestanding pronominal object, it also has a bound pronominal subject.

(10) compares the New Guinea pidgin Hiri Motu with its lexifier Motu, an Austronesian language of New Guinea. Hiri Motu (formerly known as 'Police Motu') has both Central and Non-central dialects, the Central dialect showing more features of the Motu language whose speakers are in close promixity (Dutton 1996). In particular, Central Hiri Motu preserves both an optional bound pronominal object and a bound possessive pronoun, similar to Motu, while Non-central Hiri Motu replaces these with analytic pronouns. (10) illustrates the object pronominals (Foley 1986: 33–35):

(10)

Non-central and Central Hiri Motu: 'I see you'	Central Hiri Motu:	Motu:
lau itaia oi	*lau ita-mu*	*na ita-mu*
I see you	I see-you	I see-you

The object pronoun of Non-central Hiri Motu *oi* can appear initially as well as finally.

The universalist hypothesis (already qualified in Mühlhäusler 1986; see also Mühlhäusler and Harré 1990) is rebutted by Thomason and Kaufman (1988: ch. 7), who show that pidgins may contain highly marked structures in their phonology, morphology, and syntax; see also Bakker 1995, Foley 1988, n.d., and Thomason (ed.) 1996. The bound pronominal subjects in Broken Oghibbeway, and bound pronominal objects and possessors in Central Hiri Motu are examples of such marked morphosyntactic structures.

There is also a conceptual problem faced by universalist approaches to pidgin genesis: to explicate how it is that universals (whether represented by default parameter settings or unmarked structures) enter into pidginization. The frequently repeated aphorism in (11) raises the question, How do universalist characterizations of the initial state in language learning apply to pidgin genesis at all?

(11) "Pidgin languages by definition have no native speakers ...." — Mühlhäusler (1986: p. 5)

The creators of pidgins are adult speakers of the contact languages who have already acquired fully elaborated vernacular languages. In creating the pidgin they need never be in the initial state of the language learner. As Thomason and Kaufman (1985: 172–173) argue, pidgin genesis cannot always be modelled as acquisition of a target language by a learner given restricted input (the plantation pidgin model). Sometimes there is only a process of negotiating a compromise language for restricted purposes of communication between groups of speakers of different languages, none of which is in any sense a "target language".

Pidgin genesis, according to Thomason and Kaufman (1988: 174ff) and Thomason (1997), is a result of mutual linguistic accommodation among speakers of different languages for restricted communicative purposes in an extended contact situation. Pidginization begins with speakers simplifying the structures of their native languages in order to be understood by their interlocutors. The actual degree of simplification by the developers of a pidgin, according to Thomason and Kaufman (1988: 192), depends on "the degree to which marked features of languages that they already know are shared with marked features of their interlocutors' languages" (as well as on social factors such as relative prestige or power). In cases where the interlocutors' languages are typologically disparate, the simplification process would eliminate most marked features, leaving only universally unmarked features. In cases where there is little typological distance among the languages whose native speakers are developing the pidgin, more of the shared marked features may be retained in the pidgin because they remain relatively easy to learn for the interlocutors (cf. Mufwene 1991, Thomason and Kaufman 1998: 256ff). This accommodation/unmarkedness theory of pidgin genesis leads us to Hypothesis IV:

(12) Hypothesis IV: Free pronouns are prevalent in pidgins because pidgin genesis begins with a process of simplification in which speakers accommodate their interlocutors by eliminating marked types of forms from their language which are not shared by their interlocutors' language. Free pronouns are simpler (less marked) than bound pronouns. However, pidgins arising from typologically close contact languages sharing many marked structures may retain bound pronouns.

Hypothesis IV can explain the presence of the bound pronominal subject in Broken Oghibbeway. This pidgin was used in the early nineteenth century by several Indian tribes in Wisconsin in their dealings with traders and people of mixed blood. Ojibwe has a bound pronominal system, along with complex inverse verbal morphology (Schwartz and Dunnigan 1986). In Broken Oghibbeway the inverse system is greatly simplified; a pronominal object is expressed by an independent pronoun rather than an affix, but a pronominal subject is still expressed by the verb morphology (9). This is a shared feature of the Indian source languages (Nichols 1995). Likewise, the differences between the Central and Non-central dialects of Hiri Motu also appear to reflect the typological distance between the languages of the surrounding speaker populations: the Central dialect, surrounded by languages related to Motu, shares more features of Motu; the Non-central dialect, surrounded by languages unrelated to Motu, shares fewer features of Motu.[3] Finally, Yimas Pidgin is based on Papuan contact languages such as Yimas, Arafundi, and Alamblak (Williams 1993). Though these languages are typologically similar in having bound pronominal systems, Foley n.d. shows that Yimas bound pronouns are prefixed to the verb stem, while Arafundi bound pronouns are suffixed. This morphological difference could pose an analytic difficulty for comprehension in interlingual communication, which the pidgin avoids with freestanding pronouns.

Hypothesis IV is intuitively plausible and empirically the most comprehensive of the various solutions we have examined, combining aspects of all of the previous Hypotheses. But it rests on a theory of simplification which is not provided. It simply takes knowledge of how to simplify one's language by eliminating marked structures to be a necessary precondition for pidgin genesis.Yet in most current linguistic theories the grammar of a language is a tightly interconnected system specified with an elaborate network of formal dependencies referencing hidden structure and covert categories. How is it possible formally to target a specific marked structure for elimination? A related question is, How can marked structures be distinguished from universally unmarked structures in the adult grammar? Relative markedness of structures is revealed by asymmetries found in their frequencies of occurrence across languages (Greenberg 1966). How can such knowledge be accessed in the grammar of an individual adult under this model? I will show that these questions can be answered within the framework of Optimality Theory (OT) (Prince and Smolensky 1993), and specifically within the OT-LFG framework which embeds the LFG theory of structures within the OT theory of constraint interaction (see Bresnan 1997a,b, 1998a,b and references cited therein).

The structure of pronouns in LFG

In LFG, morphological and syntactic pronominal forms lexically specify the same kinds of pronominal content (Bresnan and Mchombo 1986, 1987, Demuth and Johnson 1989, Andrews 1990, Uyechi 1991, Börjars, Vincent and Chapman 1997, Toivonen 1996, 1997, 1998, Nordlinger 1998, Bresnan forthcoming, inter alia). Thus in LFG (as in typological/functional theories) elements which function as definite personal pronouns are not structurally uniform across languages,

[3] In principle, either preservation of shared features of the source languages at the beginning of pidgin genesis, as Thomason and Kaufman 1988 theorize, or subsequent influence of the source languages after stabilization of the pidgin could account for these facts. See Siegel 1997 for recent discussion.

but show formal variation, as schematized in (13) from Bresnan 1997b:[4]

(13) **Range of personal pronominal forms:**

Zero Bound Clitic Weak Free

Zero: pronominals having no expression in morphology or syntax

Bound: morphologically bound pronominals, also called pronominal inflections, which are expressed by affixal structure on a head

Clitic: pronominals that have a specialized syntactic position and are phonologically bound to a host ('special clitics' in Zwicky's (1977, 1985) sense)

Weak: freestanding pronominal forms, neither phonologically nor morphologically bound to another constituent, but atonic and differing in syntactic distribution from Free pronouns

Free: freestanding pronominal forms which may receive primary sentence accents

Pronouns are universally characterized through their referential role and functions (represented in f-structure), not through their syntactic category and phrase structure configuration (represented in c-structure). Some elements which resemble clitic pronouns, such as the indirect object clitic copies in Spanish, are not pronominal in content, but simply markers of grammatical agreement (Bresnan and Mchombo 1987, Andrews 1990). They occur with every kind of indirect object, including negative indefinites, interrogatives, etc. Likewise, some elements which resemble non-pronouns, such as a number of Amerind personal pronouns morphologically derived from inflected verbal roots (Nichols and Peterson 1996, Lipkind 1945) or deitics used anaphorically in many languages or even an obligatory subject agreement prefix, may actually function as pronominals (Greenberg 1986: xix; Bresnan 1997b; Demuth and Johnson 1989). The major f-structure types of pronominal attributes are schematized in (14). Not all pronouns need have all properties.

(14) **Crosslinguistic properties of personal pronouns:**

PRO — variable referentiality, lack of descriptive content
TOP — topic-anaphoricity (Givón 1976, 1983, 1984, 1990: 916ff)
AGR — classification by person, number, gender (Givón 1984: 354–5)

TOP, PRO, and AGR represent types of attribute-value pairs in the f-structure of pronouns, as illustrated in (15):

(15) $\text{PRO}: \begin{bmatrix} \text{PRED} & \text{'PRO'} \end{bmatrix}$

$\text{AGR}: \begin{bmatrix} \text{PERS} & \text{I} \\ \text{NUM} & \text{SG} \end{bmatrix}$, etc.

$\text{TOP}: \begin{bmatrix} \text{NEW} & - \end{bmatrix}$ (Choi 1996)

Pronouns can be represented independently of their forms of expression by using f-structures based on these properties, as illustrated in part by (16):

[4] The characterization of the theory of pronominal markedness in this section and the next closely follows Bresnan 1997b.

(16) **Representation of pronominal content by f-structures:**

$$\begin{bmatrix} \text{TOP} \\ \text{PRO} \\ \text{AGR} \end{bmatrix} \begin{bmatrix} \text{PRO} \\ \text{AGR} \end{bmatrix} \begin{bmatrix} \text{TOP} \\ \text{PRO} \end{bmatrix} \ldots$$

The leftmost feature structure in (16) specifies a pronominal which is specialized for topic-anaphoricity and is also classified for person, number or gender. The rightmost feature structure specifies a specialized topic-anaphoric pronominal which lacks any agreement classifications.

The pronominal inventory of a language is a set of pairings of instances of structural types from (13) with feature structures representing pronominal content as in (16) (Bresnan 1997a,b). For example, in (17) bound and free pronouns are represented as the pairings of a morphological affix *af* or a syntactic category X^0, respectively, with a feature structure representing their pronominal content; and the zero pronoun may be represented as the pairing of pronominal content with no structural expression at all, either morphological or syntactic. Null structure is the absence of structure, represented by Ø.

(17) **Representation of pronominals as form/content pairings:**

Zero: $< \text{Ø}, \begin{bmatrix} \text{PRO} \\ \text{TOP} \end{bmatrix} >$ Bound: $< \textit{af}, \begin{bmatrix} \text{TOP} \\ \text{PRO} \\ \text{AGR} \end{bmatrix} >$

Free: $< X^0, \begin{bmatrix} \text{PRO} \\ \text{AGR} \end{bmatrix} >$

Just as the features PRO, AGR, and TOP in (14) abbreviate more fine-grained attributes (15), so the pronominal forms in (13) are general types, each standing for more fine-grained characterizations, such as proclitic/enclitic, prefix/suffix, and the like. We will abstract away from these complexities.

As it is, this theory of pronominal structure is rather unconstrained, because it allows the cross-product of surface forms and abstract featural content. In fact, however, we wish to limit our consideration to formally marked featural contrasts. That is, we are considering only overtly marked agreement and topicality contrasts, not abstract features that bear no relation at all to surface forms of expression. To exclude excessively abstract uses of agreement features and functional content, we will make use of the classification in (18):

(18) **Classification of pronominal forms:**

$$\underbrace{\text{Zero}}_{\textit{nonovert}} \quad \overbrace{\text{Bound} \quad \text{Clitic} \quad \text{Weak} \quad \text{Free}}^{\textit{overt}}$$

$$\underbrace{\text{Zero} \quad \text{Bound} \quad \text{Clitic} \quad \text{Weak}}_{\textit{reduced}} \quad \overbrace{\text{Free}}^{\textit{nonreduced}}$$

Overt pronominals are those with perceptible morphological or syntactic exponents; only the Zero pronominal lacks a perceptible exponent and is categorized as nonovert this sense. Reduced pronominals are those whose exponents have less phonological or morphological substance than nonreduced pronouns; only the Free pronoun is unreduced. Now we suppose that AGR features by definition occur only with overtly marked contrasts in the agreement categories.[5] We also suppose that

[5]See Bresnan 1997b for discussion of apparent counterexamples arising from the paradigmatic use of Zero pronominals.

the category of topicality in pronominals is marked by reduced expression (Givón 1983).[6] These assumptions are spelled out in (19):

(19) a. Overt ⇔ AGR: Pronominals are inherently specified for person/number/gender contrasts if and only if they are overt.

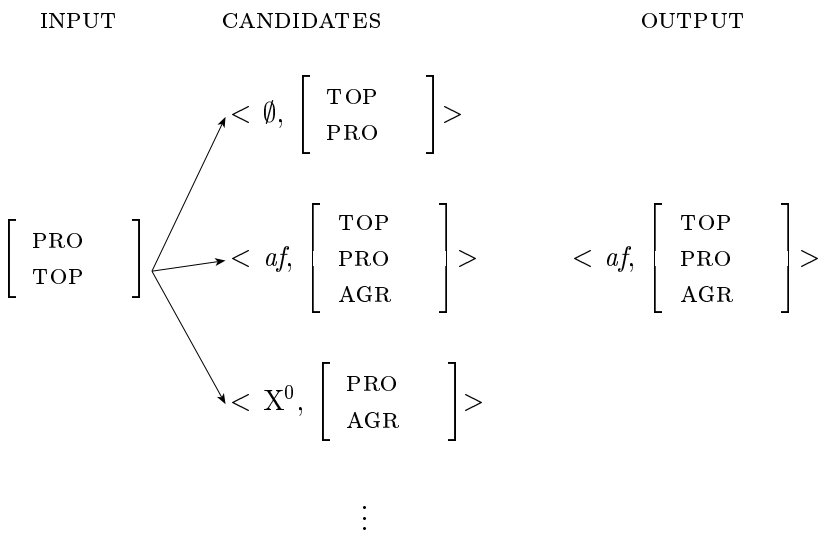
b. Reduced ⇔ TOP: Pronominals are reduced if and only if they are specialized for topic anaphoricity.

It follows that "[n]o language has an overt definite personal pronoun devoid of any distinctions of person, number, or gender, while many languages have zero pronouns with just this property" and that "[n]o language has zero, bound, or clitic personal pronouns used only for emphasis and focus, though many languages have free pronouns with this property" (Bresnan 1997b).

The markedness of pronouns in OT

Now we can take LFG to generate the typological space of possible pronominal structures. The inventory of pronominals of each language is selected from this space by optimization. For each possible specification of pronominal content (the 'input'), the grammar of a particular language will select the optimal expression of that content (the 'output') by evaluating all possible pairings of pronominal feature structures with instances of structural types, including the examples in (17), among many others. This is illustrated in (20):

(20) **OT-LFG model (Bresnan 1997a,b,c):**



Crucially, the optimal candidate need not be a perfect analysis of the input; it may overparse or underparse the input pronominal content, as illustrated in (20), where the optimal pronominal contains agreement features not specified in the input. This is possible because the evaluation function is based on a set of universal constraints which may conflict with each other. Each

[6] Of course, definite noun phrases and other referring expressions can refer to topical material (see Choi 1996 for a development within the OT-LFG framework). We are assuming, however, that reduced pronominals grammaticalize this discourse property.

language-particular grammar resolves the conflicts by a prioritization of the constraints. Thus perfect fidelity to some specified content may be overridden by a more highly ranked constraint that penalizes some other aspect of competing structures having more fidelity than the selected candidate. The selection of language-particular forms minimizes the violations of constraints according to the defintion in (21):

(21) EVAL: Given a universal set CON of possibly conflicting constraints, and given a language-particular strict dominance ranking of CON, the optimal/most harmonic/least marked candidate (= the output for a given input) is one that best satisfies the top ranked constraint on which it differs from its competitors (Grimshaw 1997, Smolensky 1996b).

In considering pronominal markedness, it important to recognize the two distinct conceptions of markedness given in (22):

(22) **Two concepts of the unmarked:**

neutral: the forms used under neutralization of oppositions (Jakobson 1931): *unmarkiert*, 'functionally unmarked' (Aikhenvald and Dixon 1998)

reduced expression: the forms having fewer morphemes or less phonological content (Jakobson 1984 [1939], Greenberg 1966, Comrie 1986): *merkmallos*, 'formally unmarked' (Aikhenvald and Dixon 1998)

Interestingly, these two concepts of the unmarked are opposed in the category of pronouns. Reduced pronouns are formally unmarked (by definition) and they are universally the preferred forms for expressing familiar, topical referents (Givón 1976, 1983, 1984, 1990; Haiman 1985: 150, 167, 194, 232–2). But free pronouns are the forms used to fill in the gaps in systems of reduced pronominals: where reduced pronominals are unavailable, the free pronoun may take on their functions (Bresnan 1997a,b). In that sense, free pronouns are functionally unmarked.

Bearing this distinction in mind, let us turn to the markedness constraints shown in (23). These are taken from Haiman 1985 via Bresnan 1997b:

(23) **Markedness constraints on pronominals:**

$*\emptyset$	$*af$	$*$ CL	$*$ Weak
[PRO]	[PRO]	[PRO]	[PRO]
Iconicity		*Avoid Allotaxy*	

Haiman's 1985 idea is that Zero and Bound pronominals violate a syntactic/semantic iconicity constraint, because they yield a non-isomorphic mapping between syntactic constituents and semantic referents and relations—the zero pronoun because it has no constituent structure at all, and the affixal pronoun because it is morphologically part of another constituent (the head) and so non-iconically maps a relation and referential role, two distinct semantic constituents, onto a single syntactic constituent. In contrast, the clitic and weak pronouns do not suffer from this defect, because by definition they are syntactic elements that are only prosodically dependent or defective. But clitics and weak pronouns have a different marked property: they are nonuniform in their syntactic distribution with free (neutral) pronouns (13). In French, for example, clitic pronouns generally appear preverbally, while free pronouns are postverbal. In West Flemish and Swedish, weak pronouns are attracted to positions (such as that of the complementizer or finite verb) from which free pronouns are excluded (Haegeman 1996, Sells 1998). This nonuniformity in syntactic

expressions of the same semantic roles or grammatical functions is called *allotaxy* by Haiman (1985: 162). Haiman observes that the avoidance of allotaxy is—along with iconicity—a major source of syntactic regularities seen in pidgins. (Foley 1988 and Mufwene 1991 also argue that iconicity, expressed as "invariance" or "semantic transparency," is important in the development of pidgins.)

By recasting the functional motivations of 'Iconicity' and 'Avoid Allotaxy' as the markedness constraints (23), we can derive Haiman's (1985) markedness explanation for pidgin pronominal systems from the initial ranking of markedness constraints above faithfulness in Optimality Theory (see Smolensky (1996a,b) for a recent exposition and further references). This initial structuring of constraints is proposed as a way to explain the acquisition of phonologies consisting of unmarked structures. If unmarked structures incur no marks, they provide no evidence for any particular constraint ranking in OT, and so will not lead to convergence on a single grammar. The solution is to hypothesize an initial state of the language learner in which markedness constraints outrank faithfulness constraints.

Accordingly, we have the two families of constraints initially ranked as shown in (24), where STRUCT refers to the markedness constraints (23) which penalize reduced pronominal structures, and FAITH designates a family of constraints which penalize pronominal forms whose feature structures do not match the input.

(24) STRUCT $\gg$ FAITH

Which types of forms are actually found in the inventory of a language depends on the relative ranking of STRUCT and FAITH constraints. The unreduced free pronouns will not be specialized for the TOP property, and hence they will be unfaithful to an input specified for the topicality feature. But that violation will matter less, given the ranking in (24) than the violations incurred by being a marked form. The ranking of all the markedness constraints above the faithfulness constraints means that it is worse to be a reduced form (thus violating iconicity or exhibiting allotaxy) than to be unfaithful to the input. Since this is true for any input (combination of pronominal content features), the marked pronominal forms will be absent in such a language (all else being equal). Only the neutral free pronouns will occur in the inventory. Hence this ranking, by the standard OT logic of markedness, yields the systematic pronominal inventory of highly analytic languages like English and pidgins.[7] The table in (25) schematically illustrates these points for a representative sample of the candidate set. (The ranking of the constraints is indicated by their left-to-right order in the tableaux columns. '!' represents a fatal violation, which eliminates a candidate. The optimal candidate, designated by the arrow, is the one which best satisfies the top-ranked constraint on which it differs from its competitors.)

[7] We omit from consideration unsystematic occurrences, such as the omitted second person pronoun in *Going home?* in English.

(25) **Ranking yielding only the unmarked pronoun:**

	Input [PRO, TOP]	STRUCT	FAITH
	Zero: [PRO, AGR]	*!	*
	Zero: [PRO, TOP]	*!	
	Zero: [PRO]	*!	*
	Bound: [PRO, AGR]	*!	*
	Bound: [PRO, TOP]	*!	
	Bound: [PRO]	*!	*
⇒	Pronoun: [PRO, AGR]		*
	⋮		

Thus the ranking shown in (25) gives us a pronominal inventory consisting only of potentially strong pronouns; that is, syntactically free pronouns capable of being tonically accented, having morphological classification for person/number/gender, and being unspecialized for topic anaphoricity. These are universally the unmarked pronouns.

When one of the markedness constraints is demoted below faithfulness, however, the form it marks enters into the inventory; that form becomes optimal for expressing topical content, as illustrated in (26). The table has been simplified by omitting all but three representative candidate types and all but the relevant instances of STRUCT:

(26) **Ranking yielding a bound pronominal:**

	Input [PRO, TOP]	$*\emptyset$ [PRO]	FAITH	$*af$ [PRO]
	Zero: [PRO, TOP]	*!		
⇒	Bound: [PRO, TOP, AGR]			*
	Pronoun: [PRO, AGR]		*!	
	⋮			

For nontopical content, the free unreduced pronoun is optimal under the same ranking; see (27):

(27) **The unmarked pronoun under the same ranking:**

	Input [PRO]	$*\emptyset$ [PRO]	FAITH	$*af$ [PRO]
	Zero: [PRO, TOP]	*!		
	Bound: [PRO, TOP, AGR]			*!
⇒	Pronoun: [PRO, AGR]			
	⋮			

It follows that the demotion of the markedness constraint admits the corresponding reduced form into the inventory, but only for topical content; the free, unreduced pronoun remains optimal elsewhere.

Because of the OT principle that languages differ systematically *only* in their rankings of the universal constraint set, this (partial) theory makes the typological prediction that there are languages with free pronouns only and no bound pronominals, and languages with both free and bound

pronominals, but no languages having only bound pronominals and lacking free any pronouns. To the extent that this prediction is borne out, it provides evidence for our hypothesis that the free syntactic pronoun is the unmarked pronominal form (that is, the neutral, *unmarkiert*, form (22)). This result is stated by Bresnan (1997a) and by Carstairs-McCarthy (1992: 165–6):

(28) **Typological asymmetry among bound and free pronoun inventories: (Carstairs-McCarthy 1992: 165–6, Bresnan 1997a)**

Only free pronouns systematic (English)
Both free and bound pronouns systematic (Chichewa, Navajo, etc.)
Only bound pronouns systematic (none or rare)

Studies of the typology of pronominal systems (Forchheimer 1952, Wiesemann 1986) confirm that while there are many languages that lack reduced pronominal forms, languages that lack freestanding pronouns are rare and arguably absent.[8]

Depending on context, the unmarked (neutral content) form can be used either inclusively, subsuming the marked, or exclusively, in opposition to the marked. Hence "the unmarked member acts as a surrogate for the entire category" (Greenberg 1966: 61). The free pronoun has just this property, as we see from the above analysis. In languages having reduced pronouns, free pronouns appear to be specialized for focus uses (Schwartz 1986), but in syntactic contexts where the reduced pronouns are prohibited, the free pronoun may take on the discourse functions of the reduced pronouns. This is why the free pronouns can fill gaps in the bound pronominal paradigm (Bresnan 1997a,b). Its neutrality is the source of its polyfunctionality, which is latent in languages having reduced pronominals, the latter being the more faithful forms for expressing topicality.

Simplification by constraint promotion

The present theory incorporates Haiman's 1985 insight that pidgins utilize highly unmarked structures characterized by iconicity and the avoidance of allotaxy. However, it does not assume that the creator of a pidgin must start from the initial state of language learning (23) in which all markedness constraints dominate faithfulness constraints. Rather, the creators of a pidgin can work from their own grammars by simplification and accommodation, as proposed by Thomason and Kaufman 1988. Simplification can be modelled as a process in which speakers eliminate marked features of their language by promoting low-ranked markedness constraints above the corresponding faithfulness constraints. Reranking the constraints in this way has the effect of removing the marked pronominal forms from the pronominal inventory. The constraints targeted for reranking are those which mark types of forms that are not understood or not easily learned by the interlocutors because they are not in the inventory of their language. These constraints are easily identifiable because of the output-oriented nature of OT constraints, together with LFG's very surface-oriented theory of syntactic structure, in which each local piece of morphology or syntax monotonically adds information that characterizes the global f-structure.

As more and more markedness constraints are promoted above faithfulness constraints, the initial state of the language learner hypothesized by Prince and Smolensky (see Smolensky 1996a,b) is approached (24). In this state the grammar produces only maximally unmarked forms common to all languages. Note that the theory does not assume that developers of pidgins have knowledge of the relative frequencies of occurrence of structures across languages. They need only have knowledge of

[8] Languages possibly lacking free pronouns include Warumungu (Simpson and Heath 1982, Simpson to appear), Winnebago (Lipkind 1945), Straits Salish (Jelinek and Demers 1994), Thai, Burmese, and Vietnamese (Cooke 1968). See Bresnan 1997b for discussion.

their own particular grammar. Language particularity (insofar as it systematic) resides only in the ranking of the substantive universal constraints shared by all languages, which is used to optimize the structures in the typological space available to all languages. By the OT logic of markedness (Smolensky 1996b), promotion of markedness constraints above their corresponding faithfulness constraints guarantees convergence of grammars toward the maximally unmarked structures of the initial state (24).

Finally, when the contact languages are typologically very close, they will share a greater number of marked structure types, and fewer constraint promotions will be required to attain a mutually comprehended medium of communication. Hence the presence of marked pronominal structures in pidgins having typologically close source languages is also predicted.

This theory of markedness and simplification in the domain of pronouns is not meant to be the full story of how pidgins develop, stabilize, and expand. Rather, it is meant to provide a new theoretical basis for just one necessary part of the story of pidgin genesis—the part that requires adult speakers to be able to access universal markedness properties of all languages starting only from knowledge of their own vernaculars.

References:

Aikhenvald, Alexandra Y. and R.M.W. Dixon. 1998. Dependencies between grammatical systems. *Language* 74, 56–80.

Andrews, Avery. 1990. Unification and morphological blocking. *Natural Language & Linguistic Theory* 8, 507–57.

Arends, Jacques, Pieter Muysken, and Norval Smith (eds.) 1995. *Pidgins and Creoles. An Introduction.* Amsterdam: Benjamins.

Bakker, Peter. 1995. Pidgins. In Arends et al. (eds.), 25–39.

Bakker, Peter. 1997. Re: Null pronouns in pidgins. Posting to the CreoLIST, May 9, 1997. On-line, University of Stockholm: http://www.ling.su.se/Creole/CreoLIST/.

Börjars, K., Vincent, N. and Chapman, C. 1997. Paradigms, periphrases and pronominal inflection: a feature-based account. In *The Yearbook of Morphology*, ed. by Geert Booij and Jaap van Marle (eds.), 1–26.. Dordrecht: Kluwer Academic Publishers.

Bresnan, Joan. 1997a. The emergence of the unmarked pronoun: Chichewa pronominals in Optimality Theory. *BLS 23.* To appear.

Bresnan, Joan. 1997b. The emergence of the unmarked pronoun II. Paper presented at the Hopkins Optimality Theory Workshop, Inner Harbor, Baltimore, May 9–12, 1997. On-line, Stanford University: http://www-csli.stanford.edu/users/bresnan/download.html.

Bresnan, Joan. 1998a. Optimal syntax. To appear in *Optimality Theory: Phonology, Syntax, and Acquisition*, ed. by Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer. Oxford: Oxford University Press. Draft available on-line, Stanford University: http://www-csli.stanford.edu/users/bresnan/download.html.

Bresnan, Joan. 1998b. Explaining morphosyntactic competition. To appear in *Handbook of Contemporary Syntactic Theory*, ed. by Mark Baltin and Chris Collins. Oxford: Blackwell. Draft available on-line, Stanford University: http://www-csli.stanford.edu/users/bresnan/download.html.

Bresnan, Joan. Forthcoming. *Lexical-Functional Syntax.* Oxford: Blackwell.

Bresnan, Joan and Sam A. Mchombo. 1986. Grammatical and anaphoric agreement. *CLS 22:2. Papers from the Parasession on Pragmatics and Grammatical Theory*, 278–97. Chicago: Chicago Linguistic Society.

Bresnan, Joan and Sam A. Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63.741–82.

Byrne, Francis and John Holm (eds.). 1993. *Atlantic Meets Pacific. A Global view of Pidginization and Creolization. (Selected Papers from the Society for Pidgin and Creole Linguistics).* Amsterdam: Benjamins.

Carstairs-McCarthy, Andrew. 1992. *Current Morphology.* London: Routledge.

Choi, Hye-Won. 1996. *Optimizing Structure in Context: Scrambling and Information Structure.* Stanford, CA: Stanford University Ph.D. dissertation.

Cooke, Joseph R. 1968. *Pronominal Reference in Thai, Burmese, and Vietnamese.* Berkeley: University of California Press.

DeGraff, Michel. 1992. Is Haitian Creole a pro-drop language? In Lefebvre (ed.), 1–21.

Demuth, Katherine and Mark Johnson. 1989. Interaction between discourse functions and agreement in Setawana. *Journal of African Languages and Linguistics* 11, 21–35.

Deprez, Viviane. 1992. Is Haitian Creole really a pro-drop language? In Lefebvre (ed.), 23–40.

Drechsel, Emanuel. 1997. *Mobilian Jargon. Linguistic and Sociohistorical Aspects of a Native American Pidgin.* Oxford: Clarendon Press.

Dutton, Tom. 1996. Hiri Motu. In Thomason (ed.), 9–41.

Foley, William A. 1986. *The Papuan Languages of New Guinea.* Cambridge: Cambridge University Press.

Foley, William A. 1988. Language birth: the processes of pidginization and creolization. *Linguistics: The Cambridge Survey. Vol. IV. Language: The Socio-cultural Context*, ed. by Frederick J. Newmeyer, 162–183. Cambridge: Cambridge University Press.

Foley, William A. n.d. Pidginization: the shadow of Universal Grammar. Sydney: University of Sydney Department of Linguistics MS.

Forchheimer, Paul. 1953. *The Category of Person in Language.* Berlin: de Gruyter.

Givón, Talmy. 1976. Topic, pronoun and agreement. In *Subject and Topic*, ed. by Charles N. Li, 149–88. New York: Academic Press.

Givón, Talmy. 1983. Topic Continuity in Discourse: An Introduction. Topic Continuity in Discourse: A Quantitative Cross-Language Study, ed. by T. Givón, 5-41. Amsterdam: John Benjamins Publishing Company.

Givón, Talmy. 1984, 1990. *Syntax. A Functional-Typological Introduction.* Vols. I and II. Amsterdam: John Benjamins Publishing Company.

Givón, Talmy. 1995. *Functionalism and Grammar.* Amsterdam: Benjamins.

Goddard, Ives. 1996. Pidgin Delaware. In Thomason (ed.), 43–98.

Greenberg, Joseph. 1966. *Language Universals. With Special Reference to Feature Hierarchies.* The Hague: Mouton.

Greenberg, Joseph. 1986. Introduction: some reflections on pronominal systems. In Wiesemann (ed), xvii–xxi.

Grimshaw, J. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28, 373–422.

Haegeman, Liliane. 1996. Object clitics in West Flemish. *Approaching Second. Second Position Clitics and Related Phenomena*, ed. by Aaron L. Halpern and Arnold M. Zwicky, 135–64. Stanford: CSLI Publications.

Haiman, John. 1985. *Natural Syntax. Iconicity and Erosion.* Cambridge: Cambridge University Press.

Heine, Berndt. 1975. Some generalizations on African-based pidgins. Paper presented at the International Conference on Pidgins and Creoles, Honolulu.

Holm, John. 1989. *Pidgins and Creoles. Vol. II. Reference Survey.* Cambridge: Cambridge University Press.

Huttar, George L. and Frank J. Velantie. 1996. Ndyuka-Trio Pidgin. In Thomason (ed.), 99–124.

Jakobson, Roman. 1984 [1939]. The zero sign. In *Russian and Slavic Grammar: Studies, 1931–1981*, 151–160. Berlin: Mouton.

Jahr, Ernst Håkon and Ingvild Broch (eds.) 1996. *Language Contact in the Arctic. Northern Pidgins and Contact Languages.* Berlin: Mouton de Gruyter.

Jelinek, Eloise and Richard A. Demers. 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70, 697–736.

Kay, Paul and Gillian Sankoff. 1974. A language-universals approach to pidgins and creoles. In *Pidgins and Creoles: Current Trends and Prospects*, 46–84. Washington, D.C.: Georgetown University Press.

Lefebvre, Claire (ed.) 1992. *Haitian and Null Subject Languages. Travaux de recherche sur le créole haïtien.* No. 11. Montreal: Groupe de recherce sur le créole haïtien, Départment de linguistique, Université du Québec á Montréal.

Lipkind, William. 1945. *Winnebago Grammar.* Morningside Heights, NY: King's Crown Press.

McConvell, Patrick. 1997a. MLF and free/bound pronouns. Posting to the CreoLIST, May 28, 1997. On-line, University of Stockholm: http://www.ling.su.se/Creole/CreoLIST/.

McConvell, Patrick. 1997b. Yimas pidgin: addendum. Posting to the CreoLIST, May 29, 1997. On-line, University of Stockholm: http://www.ling.su.se/Creole/CreoLIST/.

Mufwene, Salikoko. 1991. Pidgins, creoles, typology, and markedness. In *Development and Structures of Creole Languages. Essays in Honor of Derek Bickerton*, 123–43. Amsterdam: Benjamins.

Mufwene, Salikoko. 1996. Kituba. In Thomason (ed.), 173–208.

Mühlhäusler, Peter. 1986. *Pidgin and Creole Linguistics.* Oxford: Basil Blackwell.

Mühlhäusler, Peter and Ron Harré. 1990. *Pronouns and People: The Linguistic Construction of Social and Personal Identity.* Oxford: Basil Blackwell.

Nichols, Johanna and David A. Peterson. 1996. The Amerind personal pronouns. *Language* 72, 336–71.

Nichols, John D. 1995. The Ojibwe verb in "Broken Oghibbeway". *Amsterdam Creole Studies* 12, 1–18.

Nordlinger, Rachel. 1998. *Constructive Case: Evidence from Australia.* Stanford: CSLI Publications.

Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar.* RuCCS Technical Report #2. Piscateway, NJ: Rutgers University Center for Cognitive Science.

Roberts, Julian M. 1995. A structural sketch of Pidgin Hawaiian. *Amsterdam Creole Studies* 12, 107–23.

Samarin, William J. 1996. Arctic origin and domestic development of Chinook Jargon. In Jahr and Broch (eds.), 321–339.

Schwartz, Linda J. 1986. The function of free pronouns. In Wiesemann (ed), 405–36.

Schwartz, Linda J. and Timothy Dunnigan. 1986. Pronouns and pronominal categories in Southwestern Ojibwe. In Wiesemann (ed), 285–322.

Sells, Peter. 1998. Scandinavian clause structure and object shift. *Proceedings of LFG98*, University of Queensland, June 30–July 2, 1998.

Siegel, Jeff. 1997. Mixing, leveling, and pidgin/creole development. In Spears and Winford (eds), 111–149.

Simpson, Jane. To appear. Some features of Warumungu morphology. *Handbook of Morphology*, ed. by Andrew Spencer and Arnold Zwicky. Oxford: Blackwell.

Simpson, Jane and Jeffrey Heath. 1982. *Warumungu Sketch Grammar.* Canberra: AITSIS MS.

Smolensky, Paul. 1996a. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27.720–731.

Smolensky, Paul. 1996b. The initial state and 'richness of the base' in Optimality Theory. *Technical Report* JHU-CogSci-96-4, Department of Cognitive Science, Johns Hopkins University.

Spears, Arthur K. and Donald Winford. 1997. *The Structure and Status of Pidgins and Creoles.* Amsterdam: John Benjamins.

Thomason, Sarah G. (ed.) 1996. *Contact Languages: A Wider Perspective.* Amsterdam: John Benjamins.

Thomason, Sarah Grey. 1997. A typology of contact languages. In Spears and Winford (eds), 71–88.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics.* Berkeley: University of California Press.

Toivonen, Ida. 1996. Finnish possessive suffixes in Lexical-Functional Grammar. *Proceedings of the First LFG Conference, Rank Xerox, Grenoble, France, August 26–28, 1996.* Online, Stanford University: http://www-csli.stanford.edu/publications/LFG/lfg1.html.

Toivonen, Ida. 1997. The possessive suffixes in Finnish. *CLS 23:1.* To appear.

Toivonen, Ida. 1998. The morphosyntax of Finnish possessives. Stanford: Stanford University Department of Linguistics MS. On-line, Concordia: http://www-cmll.concordia.ca/linguistics/toivonen/idapaps.html.

Uyechi, Linda. 1991. The functional structure of the Navajo third person alternation. *CLS 27:1*, 434–46.

Van der Voort, Hein. 1996. Eskimo pidgin in West Greenland. In Jahr and Broch (eds.), 157–258.

Wiesemann, Ursula (ed). 1986. *Pronominal Systems.* Tübingen: Gunter Narr Verlag.

Williams, Jeffrey P. 1993. Documenting the Papuan-based pidgins of insular New Guinea. In Byrne and Holm (eds), 355–67.

Zwicky, Arnold. 1977. *On Clitics.* Bloomington, Indiana: Indiana University Linguistics Club.

Zwicky, Arnold. 1985. Clitics and particles. *Language* 71.283–305.