Pidgin Genesis and Optimality Theory[1]

Joan Bresnan

Department of Linguistics
Stanford University
Stanford, CA. 94305 United States

bresnan@stanford.edu

To appear in

Processes of Language Contact:
Case Studies from Australia and the Pacific

Jeff Siegel (Editor)

Montréal: Les éditions Fides.

Revision of April 17, 2000

[1] This is a revised and expanded version of Bresnan (1998c), superseding that work. I am grateful to Jeff Siegel, an anonymous reviewer, Sarah Roberts, and Miriam Meyerhoff for valuable questions and comments on the previous version. I also thank Elizabeth Traugott, Bill Foley, John Rickford, Aaron Broadwell, 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 basic means of communication between groups of adult speakers of different vernacular languages (Thomason 1997). In one widely popularized view (Bickerton 1981, Pinker 1994) pidgins are heavily depreciated as linguistic systems because of their variance across speakers and lack of consistency. Yet pidgins generally undergo a historical process of stabilization which belies this stereotype (Mühlhäusler 1986, Siegel 1986), and a number of the characteristic properties of creoles are shared with the pidgins from which they directly descend historically (McWhorter to appear). Crosslinguistically, 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 1995). 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 that might account for the prevalence of free pronouns is that pidgins simply derive their pronominal forms from their source languages. A version of this hypothesis is stated in (2):

(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]

(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, English, French
Mobilian Jargon (late 17th to 20th c.)	Choctaw, Chickasaw (Muskogean)	S.E. Amerindian
Hiri Motu	Motu (Austronesian)	Papuan, Austronesian, English, Melanesian Pidgin, various European
Eskimo Pidgin	W. Greenlandic	Danish, English, French, Russian
Chinook Jargon (19th and 20th c.)	Chinook, Nootka (Chinookan, Wakashan)	N.W. Coast Amerindian, English, French
Pidgin Delaware (17th c.)	Unami Delaware (E. Algonquian)	Dutch
Pidgin Hawaiian (late 18th to early 20th c.)	Hawaiian	English, Portuguese, German, Cantonese

[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.

All of these pidgins have free pronouns.

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

(4)	Mobilian Jargon:			Choctaw:	
	'I want water./I am thirsty.'			'I am thirsty.'	
	oka	*eno*	*banna*	*oka*	*sa-banna-h*
	water	I	want	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'			Hawaiian:		
	kela	*papale*	*oe*	*ka*	*pa:pale*	*a:u*
	DEF	hat	you	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 (strong[3]) free pronouns appear to be specialized for focus uses (Schwartz 1986, Bresnan and Mchombo 1987, and references). 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'				Yimas:
	Ama	*min*	*namban*	*kratiki-nan.*	*Na-ka-tupul.*
	1SG	3SG	toward	hit-NONFUT	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 NEAR DISTAL DEICTIC). In relation

[3]Some languages have 'weak' freestanding pronouns which function like bound pronouns. See below.

to the bound pronominal forms, these free forms are 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 deictic".

While Hypothesis I is can thus be rejected for not explaining the difference in uses of free pronouns in pidgins 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 also observes); see Foley (1988). 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. 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, Haiman 1985, Mühlhäusler 1986), 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 "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 Oghibbeway 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 form of the verb 'see' *itaia* in Non-central Hiri Motu may a fusion of the verb stem *ita-* with the Motu third person singular object suffix *-ia* (Jeff Siegel, personal communication, February 10, 1999). In the Non-central Hiri Motu dialects (and optionally in Central Hiri Motu), it is an invariable verb ending replacing the Motu object agreement morphology; the verb-final *-a* is described as a "transitivity marker" by Foley (1986: 35).

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 (typologically unusual) structures in their phonology, morphology, and syntax; see also Bakker (1995), Foley (1988), 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: 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). A related point is made by Mühlhäusler (1986: 115): "universal solutions become necessary mainly in heterogeneous communities"; he emphasizes that unstable pre-pidgin jargons are subject to individual communication strategies, including transfer of features from the substrate language, so that for relatively homogeneous groups, such as Japanese plantation workers in Hawaii, non-universal strategies including transfer would seem to have been perfectly viable. 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.[4] 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, Yimas bound pronouns are prefixed to the verb stem, while Arafundi bound pronouns are suffixed (Foley n.d.). This morphological difference could pose an analytic difficulty for comprehension in interlingual communication, which the pidgin avoids with freestanding pronouns.

[4]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. Careful sociohistorical investigation is required. See Siegel (1997) for recent discussion.

It has been argued that some features of pidgins are unlikely to be simplifications by speakers of their native languages, but are instead examples of imperfect learning of a target language (Siegel 1986: 18ff). Examples of imperfect learning include the creole French word for 'water'—*lo*, which comes from "recutting" French *l'eau*, and the Bislama word for 'ear'—*sora*, a similar recutting of French *les oreilles* (Siegel 1986: 21; personal communication, February 10, 1999). Other proposed examples of imperfect learning involve fusions of morphemes. Most of the Standard Fijian inalienable nouns come into Plantation Pidgin Fijian fused with the third person singular possessive pronoun suffix *-na*: *tinana* 'mother' from *tina-na* '3 SG's mother', *ligana* 'hand' from *liga-na* '3 SG's hand' (Siegel 1986: 116). Similarly, as noted for (10), the Motu verb stem *ita* becomes in Hiri Motu *itaia*, a fusion of the third singular object marker with the Motu verb stem (and transitive suffix *-i*). At first sight it is difficult to see such examples as simplifications.

It is important to recognize that our common-sense, pre-theoretic conception of "simple" structures does not always match the results of the linguistic theory of unmarkedness. For example, the French bare nouns *eau* 'water' and *oreilles* 'ears' mentioned above begin with onsetless syllables, a marked syllable type. Across languages, the unmarked syllable type has a consonant onset followed by a vowel and no coda—CV; this syllable structure occurs in all languages, unlike the marked types. Thus in terms of syllable structure, the forms *lo* 'water' and *sora* 'ear' are less marked than the bare noun lexifier forms *eau, oreilles* and provide a more regular syllable structure. In French, liaison obligatorily syllabifies the final consonant of the preceding determiner *les* with the initial vowel of the following noun *oreilles*: le.zɔ.ʀe. This kind of phonological regrouping into unmarked syllables has led to morphological 'recuttings' in several French dialects. For example, where Standard French has *Viens-je? Vient-il?* 'Am I coming? Is he coming?', Canadian French has *Je viens-tu? Il vient-tu?*. In the Canadian dialect the historical third person enclitic pronoun *-i(l)*, which was syllabified together with the final consonant of the verb tense inflection /*t*/ by liaison, was reanalyzed as an interrogative morpheme and analogically extended across other persons, to became *-tu* in present-day Canadian French (Picard 1991). Thus the historical development of the interrogative verb suffix *-tu* in Canadian French is an example of morphological 'recutting' which also shows several reductions in markedness—CV syllable structure, neutralized person contrasts, and the SV word order of the syntactic construction it is used in—compared to the complex inversion paradigm of Standard French.

In morphosyntax in general, unmarkedness cannot be counted solely in terms of phonological length. The presence of third person *-na* on inalienable noun stems in Plantation Pidgin Fijian reflects the neutralization of the complex inflectional person/number distinctions of Standard Fijian inalienable pronominal possessors; across languages this is a typical means of reducing the markedness of expressing contrasts. Why then, if it ceases to express any person/number contrasts, is *-na* not omitted altogether in the Pidgin? It must be remembered that the contrast between alienable and inalienable types of possession is a pervasive characteristic of Austronesian syntax. In Boumaa Fijian (a dialect which is very close to 'Standard Fijian'), inalienable nouns are bound stems which require possessive pronominal inflections (Dixon 1988: 127–128). These forms are part of an elaborated series of five different syntactic possessive constructions keyed to pronominality, nominal expression type, and animacy properties of the possessed and possessor (Dixon 1988: 119–128). In Plantation Pidgin Fijian, in contrast, possessive constructions are simply juxtapositions of the possessor with the following possessed NP (Siegel 1986: 196). In this context, the addition of *-na* to inalienable noun stems can be seen to be a reduced representative of a much more complex morphosyntactic system for expressing categories of possession; at the levels of morphology and syntax, it clearly represents a reduction of markedness. As for the Hiri Motu fusion of the third person singular object marker *-ia* with verb stems, creating an invariable transitive marker, this development, too, is in accordance with markedness theory. Across languages, third person is the least marked person for objects (Silverstein 1976, Aissen 1999), so neutralization of object person distinctions to third person represents a reduction of markedness. The historical change of an object agreement marker to an invariable transitive marker is the natural result. We see this development in Boumaa Fijian, for example, where the final vowel *a* of the transitive verb suffix derives from a third person singular pronoun (Dixon 1988: 57, and references).

In sum, the kind of simplification involved in reduction of markedness is not the same as the common-sense, pre-theoretic concept of "simplification", and it should not be viewed in opposition to imperfect learning. Both simplifications by native speakers and imperfect learning by others can contribute to pidgins through the process of mutual accommodation, eliminating marked structures which are inaccessible to interlocutors, and utilizing any marked structures which are mutually accessible, in accordance with Hypothesis IV.

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.[5] 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 syntactic structures within the OT theory of constraint interaction (see Bresnan to appear a,b, 1998a,b and references cited therein).[6]

The structure and function of pronouns

Pronouns are universally characterized through their referential role and functions (represented in f(unctional)-structure in LFG), not through their syntactic category and phrase structure configuration (represented in c(onstituent)-structure in LFG). 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 (Suñer 1988, 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 deictics used anaphorically in many languages, or even an obligatory subject agreement prefix, may actually function as pronominals (Greenberg 1986: xix; Bresnan 1998d, to appear b; Demuth and Johnson 1989). This assumption is in accordance with typologically oriented work from a variety of traditions, including functional syntax (e.g. Givón 1976, 1983, 1984, 1990, 1995, Nichols 1986, Van Valin 1996), lexical functional grammar (e.g. Mohanan 1982, Simpson 1983, 1991, Kameyama 1985, Bresnan and Mchombo 1986, 1987, Andrews 1990, Austin and Bresnan 1996, Bresnan 1998d, 2000), Optimality Theoretic syntax (Grimshaw and Samek-Lodovici 1998, Samek-Lodovici 1996, Bresnan to ap-

[5] Ferguson's work on "foreigner talk" suggests that this ability is part of linguistic competence (Ferguson and DeBose 1977, Ferguson 1982), as Foley (1988: 165) notes.

[6] The exposition of the theory of pronominal markedness in the following two sections closely follows Bresnan (to appear b).

pear a,b), and some work in the Minimalist Program (Everett 1996). All of this work analyzes the variety of pronominal forms simply as alternative forms of expression of pronominal content.

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, to appear, Nordlinger 1998, Bresnan 1998d, 2000, inter alia). Thus in LFG (as in typological/functional theories) elements which function as definite personal pronouns are not structurally uniform across languages, but show formal variation, as schematized in (13) from Bresnan (to appear b):

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

Zero Bound Clitic Weak Free

Zero pronominals have no expression in morphology or syntax. An example is Japanese zero pronouns. Note that Zero does not include cases of so-called 'pro-drop' licensed by agreement morphology; the latter are analyzed not as Zero pronouns, but as pronominal inflections—'Bound' in (13), following work by Givón (1976), Jelinek (1984), Bresnan and Mchombo (1986, 1987), Börjars, Vincent and Chapman (1997), Speas (1997), Toivonen 1996, 1997, to appear, and many others.

Bound pronominals are morphologically bound forms, also called pronominal inflections, which are expressed by affixal structure on a head. In Kiswahili, for example, the prefix *ni-* 'I' in *Ninataka kusoma* 'I want to read' is analyzed as a Bound pronoun rather than as a marker agreeing with an empty first person NP (or DP) pronoun.

Clitic pronominals are phonologically bound to a host, and have a specialized syntactic position. They are 'special clitics' in Zwicky's (1977, 1985) sense; French clitics are an example. The phonologically reduced pronouns in English examples like *Go get 'em* are not special Clitics.

Weak pronouns are freestanding pronominal forms which cannot receive primary sentence accent and differ in syntactic distribution from Free pronouns. Weak pronouns fall in between Clitics and Free (strong) pronouns. Like Clitics, they generally do not bear primary sentence accent and have special syntactic positioning, but unlike Clitics and like Free (strong) pronouns, Weak pronouns are syntactically freestanding and do not require a host. The so-called shifted object pronouns of Swedish (Platzack 1998: 116) are examples of Weak pronouns. In *Han såg mig inte* 'He didn't see me', the weak pronoun *mig* appears before the negation *inte*, where lexical NPs cannot occur. In this 'shifted' position, it cannot bear contrastive stress and cannot be conjoined: **Han såg* MIG *inte* 'He didn't see ME', **Han såg dig och mig inte* 'He didn't see you and me'. Both of these possibilities are grammatical in the 'unshifted' position following the negation, where lexical NPs and strong free pronouns may occur. Thus the Weak pronoun differs from the ordinary Free pronoun in syntactic positioning. It differs from the Clitic in that it is syntactically freestanding and need not have a host. For example, the Swedish verb can be inverted in a yes/no question leaving the shifted object pronoun stranded without a verbal host: *Såg Anna dig inte?*, 'Didn't Anna see you?'. Other languages in which Weak pronouns occur include Tagalog and Chamorro.

Free pronouns are freestanding pronominal forms which may receive primary sentence accents. English pronouns are an example. English pronouns regularly undergo phonological reduction processes, but insofar as they exhibit no syntactic variation in expression, they are excluded by the present theory.

What then are the universal properties of pronouns? They are generally defineable as basic anaphoric expressions characterized by systematically shifting reference to persons within the utterance context. The referents of 'I', 'you', and 'she' shift in the sense that they vary systematically with the speaker and addressee: 'I' refers to the speaker and excludes the addressee; 'you' refers to the addressee and excludes the speaker; and 'she' refers to a third party who is neither the speaker nor the hearer. Of course, composite phrasal expressions like 'the speaker of the present utterance', 'the addressee of this utterance', 'the woman I told you about yesterday' could be argued to have the same properties of shifting reference depending upon speaker and addressee, but these are not basic expressions syntactically. Finally, anaphoricity distinguishes pronominals from basic expressions which are pure deictics, like 'this' and 'that': though pronominals often derive historically from deictics (Greenberg 1986: xix), they must have anaphoricity as a synchronic property to be functioning as personal pronouns. (An operational definition of anaphoricity is referential dependence upon a superordinate pronoun within a sentence: the second "I" in "I said that I would come" shows anaphoricity in this sense, while "that woman" in "I said that that woman would come" does not, even when that woman is in fact the speaker of this sentence.) Similarly, pronominals sometimes derive historically from common nouns (Sugamoto 1989, Cooke 1968) and from honorific nominal phrases (Mühlhäusler and Harré 1990: 136–7), but it is their referential role and function in the synchronic grammar, not their etymology, that determines their pronominality.

The major types of pronominal properties are schematized in (14):[7]

[7] A significant feature type that is omitted here is that of social level or distance.

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

PRO — shifting reference, anaphoricity
TOP — topic-anaphoricity
AGR — classification by person, number, gender

'PRO' stands for the semantic properties shared by all personal pronominals, which include shifting reference and anaphoricity as described above. 'TOP' abbreviates the information-structural functions of personal pronouns such as specialization for reference to topical elements, the 'continuing topic' in the sense of Givón (1976, 1983, 1984, 1990: 916ff). In many languages there is a distinct series of pronominal forms which is reserved for reference to the topic; in Chichewa, for example, morphologically bound pronominal forms must be used to refer to a dislocated topic (Bresnan and Mchombo 1986, 1987). 'AGR' is a cover term for the classificatory dimensions by which personal pronominals are morphologically distinguished—universally person (allowing for participant deixis and inclusion/exclusion relations among participants), often number (singular, dual, paucal, and plural), and frequently gender (classifications into kinds) (Givón 1984: 354–5); this property is abbreviated by AGR in (14). Not all pronouns have AGR and TOP features. It must be emphasized that TOP, PRO, and AGR are cover terms standing for groups of more fine-grained properties. Various choices of specific pronominal features are compatible with the general theory.

Pronouns can be represented independently of their forms of expression by using f-structures based on these properties, as illustrated in part by (15):

(15) **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 (15) 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. Exactly what combinations of pronominal properties are possible across languages is determined by a set of constraints (Bresnan to appear b), which are encapsulated in the generaliza-

Mühlhäusler and Harré (1990: 64) take the major pronominal contrasts to be (i) "'person' and the features of participant roles" and (ii) "distance and proximity (obviative and proximative) both spatial and social".

tions of (18) below. These entail that the PRO property is always accompanied by at least one of TOP or AGR.

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 to appear a,b). The angled brackets in (16) enclose the two components of the pair, the left component being an expression type from (13), and the right component being a feature structure from (15). For example, in (16) 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.

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

$$\text{Zero:} \quad < \emptyset, \begin{bmatrix} \text{PRO} \\ \text{TOP} \end{bmatrix} > \qquad \text{Bound:} \quad < \mathit{af}, \begin{bmatrix} \text{TOP} \\ \text{PRO} \\ \text{AGR} \end{bmatrix} >$$

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

Just as the features PRO, AGR, and TOP in (14) abbreviate finer-grained attributes, 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 details.

As it is, this theory of pronominal structure is rather unconstrained, because it allows all possible pairings of expressions 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 (17):

(17) **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. We also suppose that the category of topicality in pronominals is marked by reduced expression (Givón 1983).[8] These assumptions are spelled out in (18):[9]

(18) 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 to appear b).

The form-function assocations in (18a,b) are abundantly supported by typological and functional observations as well as detailed studies of particular languages. (18a), the general correlation between zero pronouns and lack of inherent specification for AGR properties, is observed by Bresnan (to appear b). It is supported by the fact that in languages which lack verbal agreement morphology, zero pronouns are not restricted as to person and number. This is true for Japanese, Chinese, Malayalam, Jiwarli (Austin and Bresnan 1996: 248–50), and many other languages.

Contrary to constraint (18a), some languages do appear at first sight to have zero pronouns specified for person, number, or gender: for example, Warlpiri has a definite third person singular zero pronominal object (Hale 1973), and Brazilian Portuguese has a definite zero pronominal object which can be used only in the third person (singular or plural) (Farrell 1990: 328). However, in these cases the zero pronouns are filling paradigmatic gaps in the bound or clitic pronominal system. Thus Warlpiri has overt bound pronominal markers for subject and object on its Auxiliary in all persons and numbers,

[8] Of course, definite noun phrases and other referring expressions can refer to topical material (see Choi 1999 for a development within the OT-LFG framework). We are assuming, however, that reduced pronominals grammaticalize this discourse property.

[9] In Bresnan (to appear b) these definitions are developed as Harmony constraints. The present version is a convenient simplification.

except for the third person singular object (Hale 1973, Nash 1996: 121), which is precisely the gap filled by the zero. Similarly, Brazilian Portuguese has a more restricted system of pronominal clitics than other Romance languages; its third person accusative forms, singular and plural, are "no longer vital" in the language (Farrell 1990: 327), and these are precisely the gaps filled by the zero. In such cases, the restriction of the zero pronoun to uses requiring specific featural content follows from morphosyntactic competition: the bound pronominal forms block the use of the zero wherever their own featural specifications apply, leaving the zero to be used elsewhere. The main point to note is that because the featural values of the zeros in these cases are predictable, it is unnecessary and unexplanatory to specify them as intrinsic properties of a null pronominal form. We neededn't postulate multiple Zeros, each lexically specified for inherent features such as third person or singular number. The paradigmatic blocking effect that yields specific featural interpretations for unmarked forms is analyzed in OT in Bresnan (1998b).[10]

(18b) refers to the inventories of grammatical pronominal types specified in (17). In languages that have both reduced and nonreduced pronouns as morphosyntactically distinct grammatical series, the two types contrast in function, with the reduced pronominals being specialized for topic-anaphoricity (Givón 1984, 1990: 917) and the nonreduced pronouns having focus functions (Schwartz 1986). The general correlation between reduced form and topic-anaphoricity is observed by Givón (1984, 1990: 917) under the name 'referential iconicity'; similar observations are made by Haiman (1985), Van Valin (1996), and many others; see Bresnan (to appear b) for further references and discussion. As we will see below, the fact that reduced pronouns are not specialized for topic-anaphoricity does not mean that they cannot be used to refer to continuing topics. Rather, it means only that they are not required to refer to the continuing topic, as reduced pronouns are.

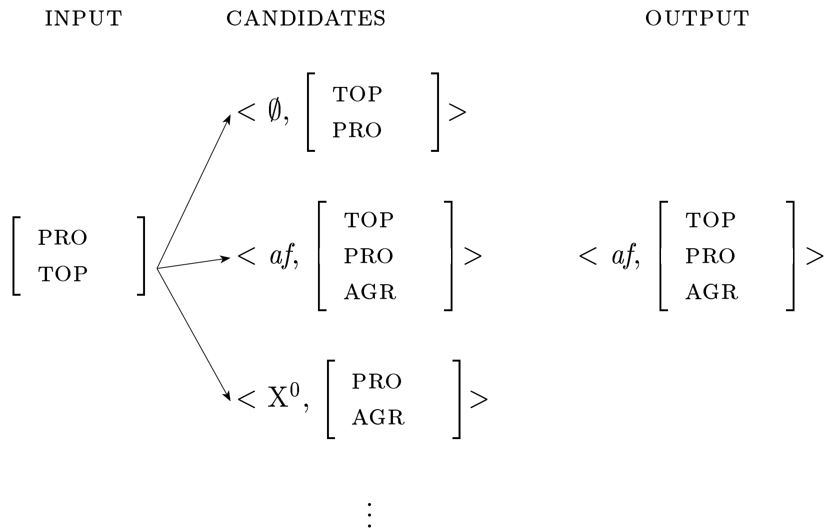
The markedness of pronouns in OT

Now we can take LFG to generate the typological space of possible pronom-

[10]There are well-known cases in English where Zeros appear to have specific agreement features: left peripheral ellipsis of the subject from a conjunct (*Mary is leaving, but __ is leaving despite herself*), and imperatives (*Be quiet, won't you?*), as an anonymous reviewer points out. Notice that in both of these cases, the interpretation of the Zero is completely determined by the context of the construction. In left peripheral ellipsis, it just matches the features of the subject of the left conjunct, and as the subject of imperatives, it must be second person. Therefore, there is no need to postulate multiple lexical Zeroes inherently specified for different agreement features in these cases. Rather, they follow from the constraints determining the construction properties.

inal 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 (16), among many others. The general form of an OT grammar is illustrated in (19):

(19) **OT-LFG model (Bresnan to appear a,b):**



The generator (GEN) is given an input. In this case, the input is some specification of possible pronominal content, such as the topical pronominal shown in (19). It could be any possible specification of pronominal content, because the input is universal, and must not vary for particular languages. In Optimality Theory systematic differences among languages—such as whether closed syllables belong to the inventory, whether verbs must be final, or whether clitic pronouns occur—are derived by constraint ranking, not by specifications of different inputs. For each particular input, GEN generates a set of possible candidate analyses. Crucially, the optimal candidate for a given input need not be a perfect analysis of that input; it may overparse or underparse the input pronominal content, as illustrated in (19), where the optimal pronominal contains agreement features not specified in the input. Varying analyses of the input across languages are possible because the evaluation function is based on a set of universal constraints which may conflict with each other.

The two major types of constraints are Faithfulness constraints (FAITH) and markedness constraints (STRUCT). Markedness constraints penalize complex or 'difficult' structures. Faithfulness constraints require that features of the input content be preserved in the output expression. They thus serve the communicative function of expressing contrasts in content, protecting content against the eroding effects of markedness constraints on forms. Obviously FAITH and STRUCT constrainst conflict. Becauase constraints are universal and are present in the grammar of every particular language, each language-particular grammar must resolve the conflicts by a prioritization of the constraints. Thus perfect faithfulness to some specified content may be overridden by a more highly ranked constraint that penalizes some other aspect of competing structures having more faithfulness than the optimal candidate. The selection of language-particular structures minimizes the violations of constraints according to the definition in (20):

(20) 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 is important to recognize the two distinct conceptions of markedness given in (21):

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

neutral: the forms used under neutralization of oppositions (Jakobson 1984 [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 to appear a,b). In that sense, free pronouns are functionally unmarked.

Bearing this distinction in mind, let us turn to the markedness constraints shown in (22). These are taken from Haiman (1985) via Bresnan (to appear b):

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

$$\underbrace{\begin{array}{cc} *\emptyset & *\mathit{af} \\ [\textsc{pro}] & [\textsc{pro}] \end{array}}_{\textit{Iconicity}} \underbrace{\begin{array}{cc} *\textsc{cl} & *\text{Weak} \\ [\textsc{pro}] & [\textsc{pro}] \end{array}}_{\textit{Avoid Allotaxy}}$$

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 object pronouns are postverbal. In Swedish, weak pronouns are attracted to positions (such as that of the complementizer or finite verb) from which free pronouns are excluded (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 (22), 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. If STRUCT ≫ FAITH, violations of structural markedness constraints are worse than violations of faithfulness constraints. Hence, being a structurally marked form will be worse than failing to preserve contrasts. In order to minimize violations, the marked forms will be avoided in favor of unmarked forms, regardless of the input (content).

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

(23) STRUCT ≫ 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 (23) 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.[11] The table in (24) 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.)

[11] We omit from consideration unsystematic occurrences, such as the omitted second person pronoun in *Going home?* in English. In genres such as 'recipe syntax' (*Take 5 eggs, beat till stiff, pour into dish*), Zero pronouns are systematically used in English; this fact can be explained by postulating a different ranking of constraints for the genre, as suggested by an anonymous reviewer.

(24) **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]	*!	*
⇒	Free: [PRO, AGR]		*
	⋮		

Thus the ranking shown in (24) 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 (25). The table has been simplified by omitting all but three representative candidate types and all but the relevant instances of STRUCT:

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

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

For nontopical content, the free unreduced pronoun is optimal under the same ranking; see (26):

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

	Input [PRO]	*∅ [PRO]	FAITH	*af [PRO]
	Zero: [PRO, TOP]	*!		
	Bound: [PRO, TOP, AGR]			*!
⇒	Free: [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 any free 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 (21)). This result is stated by Bresnan (to appear a,b) and by Carstairs-McCarthy (1992: 165–6):

(27) **Typological asymmetry among bound and free pronoun inventories:**

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

Studies of the typology of pronominal systems (Forchheimer 1953, Wiesemann 1986) confirm that while there are many languages that lack reduced pronominal forms, languages that lack freestanding pronouns are rare and arguably absent.[12]

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

[12]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 (to appear b) for discussion.

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 to appear a,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.

Markedness reduction by constraint demotion

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 (22) 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 reranking low-ranked markedness constraints above the individual faithfulness constraints that conflict with them. Knowledge of the initial state is not necessary for this process. All that is needed is the current ranking of the speaker's own grammar, and the ability to identify the conflicting markedness and faithfulenss constraints in that grammar. Reranking individual markedness constraints above the corresponding faithfulness constraints in conflict with them 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 reranked by this process above the faithfulness constraints that conflict with them, the initial state of the language learner hypothesized by Prince and Smolensky (see Smolensky 1996a,b) is approached (23). In this state the grammar produces only maximally unmarked forms common to all languages.

This 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 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), Demotion of faithfulness constraints below their corresponding markedness constraints guarantees convergence of grammars toward the maximally unmarked structures of the initial state (23).

Finally, when the contact languages are typologically very close, they will share a greater number of marked structure types, and fewer constraint demotions 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.

Let us examine in a little more detail why this theory of markedness reduction does not presuppose direct access to the initial state of the language learner by the the adult. Convergence toward the initial state from the adult state by means of constraint reranking is possible in OT because the same universal constraints are already present in every particular language. The grammar of a language having a highly marked inventory of pronominal forms has exactly the same constraints as the grammar of a language having only the unmarked forms; what differs are the relative *rankings* of constraints. As observed above, markedness of output forms can be reduced by noticing 'difficult' or unsuccessfully comprehended forms and, on the basis of the 'marks' (the patterns of constraint violations) assigned to the flagged output forms, demoting the constraints that favor them over competitors. In (25), for example, the highest ranked constraint that favors the marked Bound pronominal over the next-best competitor, Free, is FAITH. If FAITH is demoted, as in (28), the Bound pronominal is eliminated from the inventory in favor of the less marked Free pronoun:

(28) **Ranking eliminating a bound pronominal (cf. (25)):**

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

Unmarked forms are present in every language. Even when marked forms are optimal, the unmarked variants are latent in the grammar, waiting to emerge in contexts where faithfulness to input contrasts (which favors marked forms) is overriden. This 'emergence of the unmarked' effect has been documented in OT studies in phonology, and it occurs in morphosyntax as well (Bresnan to appear a,b; Lee to appear). The crucial point of the present study is that the creators of a pidgin can exploit this latent availability of unmarked forms in their own grammars. Pidgins reveal a massive emergence of the unmarked.

Several algorithms for learning OT grammars have been developed (Tesar and Smolensky 1998, Boersma and Hayes 1999), and these might be modified to give a formalization of the above model. However, the process of markedness reduction proposed here does *not* require reversing the process of first language acquisition or remembering the learner's own acquisition history. In the course of first-language acquisition, a speaker may have gone through many rerankings, moving the same constraints up and down again as various data are encountered and analyzed. In adult simplification there is no need to trace the same path in reverse, and it is highly unlikely that an adult speaker would do so, since the sequence of data encountered is different. It is the results of the markedness reduction process that will bear resemblances to the hypothesized initial state, but the sequence of processes involved—in terms of reranking—need not be the same.

In view of the convergence of pidginization on rankings yielding unmarked structures found in the initial state of first language acquistion, one might expect the syntactically unmarked pronominal forms to play a role during acquisition as well in pidgin genesis. However, we must be cautious about such inferences. As we have seen above, markedness in morphosyntax is very different from markedness in phonology, undoubtedly because the elementary units in phonology (feaatures) do not carry meanings, as the elementary units in synax (words or morphemes) do. Hence, there are many more layers of complexity in lexical and syntactic acquisition—many more dimensions of unmarkedness which may conflict. If meaning and pragmatics are more salient than the fine points of morphosyntactic structure during early acquisition, one might not see much evidence of *syntactic* markedness in this domain for a while, since the other markedness dimensions will be getting much more play. For example, Japanese freestanding pronouns are structurally unmarked according to our theory, yet they are freighted with implications of social distance and politeness that drastically constrain their use in the adult language. We might model this in OT as an overlay set of constaints which skew the markedness relations obtained purely in terms of those instances of Struct and Faith hypothesized above.

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.

Aissen, Judith. 1999. "Markedness and subject choice in Optimality Theory." *Natural Language & Linguistic Theory* 17: 673–711.

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.

Austin, Peter and Joan Bresnan. 1996. "Nonconfigurationality in Australian aboriginal languages." *Natural Language & Linguistic Theory* 14: 215–68.

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

Barbosa, Pilar, Danny Fox, Paul Hagstrom, Martha McGinnis, and David Pesetsky (eds). 1998. *Is the best good enough? Optimality and competition in syntax.* Cambridge, Massachusetts: The MIT Press and MIT Working Papers in Linguistics.

Bickerton, Derek. 1981. *The roots of language.* Ann Arbor: Karoma Publishers.

Boersma, Paul and Bruce Hayes. 1999. "Empirical tests of the Gradual Learning Algorithm." On-line, Rutgers University: Rutgers Optimality Archive, ROA-348-1099, http://ruccs.rutgers.edu/roa.html.

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 (ed). 1982. *The Mental Representation of Grammatical Relations.* Cambridge: The MIT Press.

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.

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. 1998c. "Pidgin genesis in Optimality Theory." *Proceedings of LFG98, University of Queensland, June 30–July 2, 1998*, ed. by Miriam Butt and Tracy Holloway King. On-line, CSLI Publications: http://csli-publications.stanford.edu/LFG/3/.

Bresnan, Joan. 1998d. "Morphology competes with syntax: explaining typological variation in weak crossover effects." In Barbosa et al. (eds), 59–92.

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

Bresnan, Joan. To appear b. "The emergence of the unmarked pronoun." In *Optimality-theoretic syntax*, ed. by Jane Grimshaw, Géraldine Legendre, and Sten Vikner. Cambridge, MA: The MIT Press.

Bresnan, Joan. 2000. *Lexical-functional syntax*. Oxford: Blackwell.

Bresnan, Joan and Sam A. Mchombo. 1986. "Grammatical and anaphoric agreement." *CLS 22. 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 Chichewâ." *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.

Comrie, Bernard. 1986. "Markedness, grammar, people, and the world." In Eckman, Moravscik and Wirth (eds), 85–106.

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.

Dixon, R. M. W. 1988. *A grammar of Boumaa Fijian.* Chicago: The University of Chicago Press.

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.

Eckman, Fred R., Edith A. Moravcsik, and Jessica R. Wirth (eds). 1986. *Markedness.* New York: Plenum Press.

Everett, Daniel L. 1996. *Why there are no clitics. An alternative perspective on pronominal allomorphy.* Arlington, Texas: The Summer Institute of Linguistics and The University of Texas at Arlington.

Farrell, Patrick. 1990. "Null objects in Brazilian Portuguese." *Natural Language & Linguistic Theory* 8: 325–46.

Ferguson, Charles. 1982. "Simplified registers and linguistic theory." In *Exceptional language and linguistics*, ed. by L. Obler and L. Wise, 49–66. New York: Academic Press.

Ferguson, Charles and Charles DeBose. 1977. "Simplified registers, broken language and pidginization." In *Pidgin and creole linguistics*, ed. by A. Valdman, 99–125. Bloomington: Indiana University Press.

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. 1991. *The Yimas language of New Guinea.* Stanford: Stanford 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." In *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, Jane. 1997. "Projection, heads, and optimality." *Linguistic Inquiry* 28: 373–422.

Grimshaw, Jane and Vieri Samek-Lodovici. 1998. "Optimal subjects and subject universals." In Barbosa et al. (eds), 193–219.

Haiman, John. 1985. *Natural syntax. Iconicity and erosion.* Cambridge: Cambridge University Press.

Hale, Kenneth. 1973. "Person marking in Walbiri." In *A festschrift for Morris Halle*, ed. by Stephen Anderson and Paul Kiparksy, 308–44. New York: Holt, Rinehart and Winston.

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 [1931]. "Structure of the Russian verb." *Roman Jakobson. Russian and Slavic grammar. Studies 1931–1981*, ed. by Linda R. Waugh and Morris Halle, 1–14. Berlin: Mouton Publishers.

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. 1984. "Empty categories and nonconfigurational languages." *Natural Language & Linguistic Theory* 2: 39–76.

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

Kameyama, Megumi. 1985. *Zero anaphora: The case of Japanese.* Stanford, CA: Stanford University Ph.D. dissertation.

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.

Lee, Hanjung. To appear. "Markedness and word order freezing." In Sells (ed).

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/.

McWhorter, John. To appear. "Defining 'creole' as a synchronic term." In *Papers from the degrees of restructuring conference, Regensburg*, ed. by Edgar Schneider and Ingird Neumann-Holzschuh. Amsterdam: John Benjamins.

Mohanan, K. P. 1982. "Grammatical relations and clause structure in Malayalam." In Bresnan (ed), 504–89.

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.

Naro, Anthony. 1973. "The origin of West African Pidgin." *CLS 9*, 442–9.

Nash, David. 1996. "Pronominal clitic variation in the Yapa languages." In *Studies in Kimberley languages in honour of Howard Coate*, ed. by William McGregor, 117–38. Munich: Lincom Europa.

Nichols, Johanna. 1986. "Head-marking and dependent-marking grammar." *Language* 62: 56–119.

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.

Picard, Marc. 1991. "Clitics, affixes, and the evolution of the question marker *tu* in Canadian French." *French Language Studies* 1: 179–187. Printed in Great Britain.

Pinker, Steven. 1994. *The language instinct.* New York: William Morrow and Company.

Platzack, Christer. 1998. *Svenskans inre grammatik – det minimalistiska programmet.* Lund: Studentlitteratur.

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, J. 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.

Samek-Lodovici, Vieri. 1996. *Constraints on subjects. An Optimality Theoretic analysis.* New Brunswick, NJ: Rutgers University Ph.D. dissertation.

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*, ed. by Miriam Butt and Tracy Holloway King. On-line, CSLI Publications: http://csli-publications.stanford.edu/LFG/3/.

Sells, Peter (ed). To appear. *Formal and empirical issues in Optimality Theoretic syntax.* Stanford, CA: CSLI Publications.

Siegel, Jeff. 1986. *Language contact in a plantation environment.* Cambridge: Cambridge University Press.

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

Silverstein, Michael. 1976. “Hierarachy of features and ergativity.” In *Grammatical categories in Australian languages*, ed. by R. M. W. Dixon, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Simpson, Jane. 1983. *Aspects of Warlpiri morphology and syntax.* Cambridge, MA: MIT Ph.D. dissertation.

Simpson, Jane. 1991. *Warlpiri morpho-syntax: A lexicalist approach.* Dordrecht, Holland: D. Reidel Publishing Company.

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.

Speas, Margaret. 1997. "Optimality Theory and syntax: Null pronouns and control." In *Optimality Theory. An overview*, ed. by Diana Archangeli and D. Terence Langendoen, 171–99.

Sugamoto, Nobuko. 1989. "Pronominality: A noun-pronoun continuum." In *Linguistic Categorization*, ed. by Roberta Corrigan, Fred Eckman, and Michael Noonan, 267–91. Amsterdam: John Benjamins.

Suñer, Margarita. 1988. "The role of agreement in clitic-doubled constructions." *Natural Language & Linguistic Theory* 6: 391–434.

Tesar, Bruce and Paul Smolensky. 1998. "The learnability of Optimality Theory: An algorithm and some basic complexity results." *Linguistic Inquiry* 29: 229–68.

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. To appear. "The morphosyntax of Finnish possession." *Natural Language & Linguistic Theory.*

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

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

Van Valin, Robert. 1996. "Assumptions and principles of Role and Reference Grammar." Handout. Nijmegen: Max-Planck-Institut für Psycholinguistik Conference on Future Developments in Linguistic Theory, June 28, 1996.

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.