The areal linguistics of Amazonia

Patience Epps and Lev Michael

1 Introduction

Amazonia is one of the most linguistically diverse regions of the world, with a density of distinct genetic groupings – some fifty families and isolates – rivaled only by New Guinea. We know very little about the historical processes that have shaped the Amazonian linguistic picture, or gave rise to its plethora of languages. However, despite common assumptions, the Amazon basin provides ample evidence that the maintenance of diversity does not entail a lack of contact among groups speaking different languages. In many areas of the lowlands, contact among Amazonian language communities has been intense and long-term. Such contact situations have themselves profoundly shaped the linguistic profile of Amazonia and neighboring regions, giving rise to zones of typological similarity that cross-cut genetic linguistic differences.

Working out a precise account of how contact has influenced the languages of lowland South America presents a “vast and almost intractable” problem, as observed by Muysken (2012: 235). A significant aspect of this intractability lies in the paucity of descriptive and historical work that has been carried out on these languages. For many languages, it is already too late – they have been extinguished before they could be documented. However, the past few decades have seen an explosion of high-quality descriptive work on many surviving Amazonian languages, some of which are highly endangered, with new historical work building closely upon this foundation. We have thus entered an exciting new period of investigation into Amazonian language history, which is already yielding fresh insights into linguistic areality in the Amazon basin. In this chapter, we offer our assessment of the current state of the art in understanding linguistic areality in Amazonia.

Amazonia, which we define loosely here as the lowland region drained by the Amazon and Orinoco Rivers and extending to the northern and eastern litorals of the continent (cf. Dixon and Aikhenvald 1999: 4, Rodrigues 2000: 15), is bordered by the Andes mountains to the west, the Caribbean and Atlantic oceans in the north and east, and the drier regions of the Gran Chaco to the south. Most of this vast area is covered by tropical rain forest, with pockets of savannah on the margins. The majority of the linguistic diversity encountered within the Amazon region is concentrated on the western periphery, for reasons that are currently as mysterious as those behind the overall diversity itself. The major language families that do exist, most notably Arawak, Carib, Tupí, and Macro-Jê, are characterized by predominantly non-contiguous distributions, with their members interspersed by many other smaller families and isolates (see e.g. O’Connor and Muysken 2014, Campbell and Grondona 2012 for further discussion). While this linguistic

---

1 Epps gratefully acknowledges support from the National Science Foundation (HSD-0902114).
patchwork renders the task of investigating language contact complex, it also makes it to some degree more accessible, since contact-induced change tends to be easier to identify among unrelated languages.

While our focus here is on the linguistic effects of contact, we emphasize that contact necessarily takes place among speakers, and it is the dynamics of these speakers’ interactions that produce the particular linguistic outcomes that we observe. The contact-motivated similarities that cross-cut language boundaries are evident not only in lexicon and grammar, but also in discourse and sociocultural practice more generally, and include the ways in which people tell stories, sing songs, prepare food, raise children, heal the sick, and so on. Speaker interactions are grounded in, and structured by, approaches to trade, intermarriage, and ritual and convivial practice. The extent of these interactions should not be underestimated; recent work has demonstrated that many areas of Pre-Colombian Amazonia hosted large population densities and relatively complex societies (e.g. Heckenberger and Neves 2009). Documentation of long-distance trade networks (e.g. Vidal 2000, Nordenskiöld 1922, cf. Hornborg 2005) and of migrations over large distances (e.g. Clastres 1995) likewise indicates that many native Amazonians had ample opportunity to interact with other language groups.

Our discussion is primarily concerned with the effects of contact among indigenous languages, and most of the patterns we consider are undoubtedly rooted in pre-Columbian social dynamics. However, the displacement and restructuring of many Amazonian societies in the centuries following the European invasion have introduced profound changes to those dynamics, and it is in many cases unclear to what extent contemporary socio-cultural practices represent continuity with pre-Columbian times, or how patterns of linguistic diffusion have been altered over the last five hundred years. Moreover, contact with European languages, particularly Spanish and Portuguese, as well as with European-mediated languages such as Quechua and Nheengatú (a Tupí-Guaraní language promoted by early Jesuit missionaries as a lingua franca), have profoundly affected many indigenous languages, in many instances culminating in language shift (see e.g. Muysken 2012). Interestingly, these contact scenarios are often characterized by a rather different mix of processes from those observed among indigenous Amazonian languages, with considerably more code-switching, lexical borrowing, and language shift – in keeping with the different sorts of social relations that pertain among these groups. We will not address these differences in detail here, but will focus on the observable outcomes of contact in indigenous Amazonian contexts, which have tended toward grammatical diffusion and language maintenance.

This chapter is organized as follows. We begin in Section 2 with a discussion of the principal localized contact zones that have been investigated within Amazonia (the Vaupés, the Upper Xingu, and other areas), where speakers of multiple languages live in close proximity and engage in frequent interaction. Section 3 provides a wider scope, with an assessment of evidence for larger areal diffusion zones within the lowlands. This section also addresses the possibility that the Amazon basin as a whole might represent a single large-scale diffusion zone, with substantive contrasts between it and other South American regions such as the Andes and the Southern Cone. Section 4 summarizes our current understanding of Amazonian areal linguistics and outlines directions for future research.
2 Localized diffusion within Amazonia

Most historically recent situations of regular contact among speakers of different indigenous languages are found in localized zones within the South American lowlands, and it is in these contexts that areal diffusion is most easily identified. While all of these zones have been profoundly affected by the European-derived national society, they have maintained aspects of their traditional social structures and cultural practices, suggesting a degree of continuity with pre-Colombian dynamics. Many of these contemporary contact zones share some notable similarities: in particular, they are characterized by multiple groups speaking different languages, who maintain a relatively egalitarian relationship with respect to one another, and whose interaction is frequent, conventionalized, and profound – a context that clearly favors areal diffusion. Within these ‘regional systems’, as they are sometimes termed, groups tend to be characterized by a striking degree of cultural homogeneity on one hand, but on the other by a set of locally salient differences, such that they function rather like a set of interlocking cogs in a single machine (see Epps forthcoming). Map 1 illustrates the location of the principal regions discussed here.

Language plays a recurrent role as an emblem of difference in these zones, and local ideologies of language tend to strictly constrain the mixing of codes, even where frequent interaction among groups fosters intensive multilingualism (Hill 1996, Epps forthcoming). We widely encounter long-term language maintenance, limited code-switching, and low levels of lexical borrowing, often buttressed by explicit articulations of the quality and importance of linguistic difference. At the same time, particularly where individual multilingualism is high, there is considerable convergence among languages on a grammatical level. This combination of low lexical borrowing with substantial grammatical diffusion is in striking contrast with the outcomes of language contact in many other parts of the world, where lexical borrowing tends to precede and facilitate grammatical diffusion (e.g. Thomason and Kaufman 1988); nevertheless, it appears to be commonplace in Amazonia. This outcome is undoubtedly linked to speakers’ conscious efforts to avoid language mixing – of which they are most aware on the level of lexical forms – and their frequent exposure to multiple codes, which fosters the convergence of grammatical structures and categories below speakers’ “limits of awareness”, in Silverstein’s terms (1981; see also Aikhenvald 2001a and Mithun, this volume).

In many of these interactive zones, our assessment of the effects of contact on the languages themselves is facilitated by the presence of distinct language families, and by the possibility of comparison with related languages outside the region. These factors allow us to attribute similarities within the area to contact rather than to common inheritance, especially where the shared features are too numerous and/or too unusual to be easily explained as independent innovations (see Campbell et al. 1986, Epps et al. 2013). While this historically grounded approach yields relatively robust evidence of contact-induced change, cases where related languages are undocumented or do not occur outside the region require us to fall back on the identification of a set of features shared among languages within the area. Where these features cannot be shown to be significantly different from those
that exist in the languages outside the area, the case for a regional contact zone is weakened (see Campbell and Grondona 2012, Muysken 2012); however, solid comparative analysis can provide evidence for areal diffusion regardless of whether we can define a precise contrast between a particular region and the neighboring areas.

Map 1  Lowland South American contact zones

2.1 The Vaupés region

The Vaupés region of the northwest Amazon has received the most in-depth attention of any South American contact zone. The area of the Vaupés river basin is home to dozens of languages belonging to four distinct families, Tukanoan, Arawak, Nadahup, and Kakua-Nukak (formerly lumped together with Nadahup to form the ‘Makú’ group). The Vaupés region is a particularly intensive contact zone within the larger Upper Rio Negro basin, which itself appears to be a region of less profound areal diffusion (see Aikhenvald 1999a, Epps and Stenzel 2013). While most of the existing work concerning language contact in the Vaupés has focused on particular languages or on particular linguistic features, the following discussion offers a short synthesis of our current understanding of diffusion across the region as a whole, and situates it within the wider lowland South American context.

2 The post-colonial arrival Nheengatú (Tupi-Guarani) is also marginally represented in the region, but is not discussed here.
The Vaupés has been described as a regional ‘system’, in which distinct groups function together, and as a culture area, in that these groups share many features in common (see e.g. Bruzzi 1977, Epps and Stenzel 2013). Similarities include various aspects of material culture, such as house construction, manioc-processing technology, and bodily adornment, and likewise many ritual practices. Discourse practices throughout the region are also strikingly similar, despite their delivery in different languages, with shared song traditions, stories, incantations, and conversational norms. On the other hand, certain salient differences underscore the systemic nature of the region, such that groups maintain an identity as distinct, interactive units. Among the most locally meaningful of these differences is language, which is emphasized in the context of linguistic exogamy – obligatory marriage across language groups – practiced by most of the East Tukanoan and some Arawak groups in the region. Another salient difference is subsistence pattern, with an opposition between the Nadahup and Kakua-Nukak ‘forest peoples’, who prioritize hunting and gathering, and the East Tukanoan and Arawak ‘river peoples’, who focus on fishing and farming. These subsistence distinctions are accompanied by social asymmetries (such that the river peoples are ‘ranked’ more highly than the forest peoples). Interaction among regional units is also promoted by trade specializations (for example, the Nadahup people provide meat and baskets; the East Tukanoan Tuyukas the canoes, etc.).

The systemic nature of the region engenders widespread multilingualism, coupled with language maintenance. East Tukanoan peoples tend to speak several of each other’s languages, as did Tariana (Arawak) speakers,\(^3\) facilitated by the linguistic exogamy system. The forest peoples in the region have traditionally spoken at least one river-Indian language; in their case, however, bi- or multilingualism has been almost always unidirectional, due to the local social imbalance. The active maintenance of distinct languages, despite intense multilingualism and even social asymmetry, has been viewed as an outcome of linguistic exogamy, which explicitly links marriageability to linguistic difference (Sorensen 1967, Jackson 1983, Aikhenvald 2001a, Stenzel 2005, *inter alia*). However, Epps (forthcoming) argues that the salience of the link between language and identity is probably as much a cause as an effect of linguistic exogamy, in light of broader trends across lowland South America (see below).

The Vaupés emphasis on language as an emblem of identity translates into tight constraints against the mixing of languages. Intrasentential code-switching is avoided and socially condemned (see e.g. Aikhenvald 2001a: 412, Chernela 2013: 213). East Tukanoan speakers explicitly characterize their awareness of linguistic differences, stating for example that some languages “flow slowly and smoothly, ‘like waves of water’” while others “‘sound like lightning’ … with sharp angles, stops, and starts” (Chernela 2013: 216-217). In keeping with these maintenance practices, the region’s languages have experienced remarkably low levels of lexical borrowing despite the intense, long-term contact among them. A systematic lexical study of ten different Vaupés languages reveals only 2-4% loans in basic vocabulary (Bowern et al. 2011, 2014; see also Epps 2009), with slightly higher

---

\(^3\) However, recent decades have seen Tariana and many other East Tukanoan speakers shifting to Tukano, which has gained a new, colonially mediated level of status (see Stenzel 2005, Aikhenvald 1999a: 387), and in some contexts to Portuguese.
levels in flora-fauna and material/ritual culture terms (around 10-12% in Nadahup and Kakua-Nukak languages; lower in Tukanoan and Arawak). On the other hand, the existing loans include various Wanderwörter (see Bowern et al. 2014, Epps 2012), which offer the impression that if a lexical item is to be borrowed, it is likely to travel widely, possibly because it loses an association with a particular language in the process and thus becomes ‘fair game’ (see also Muysken 2012: 252).

Despite the low levels of lexical borrowing, the effects of contact in the region’s languages have had significant effects on local lexicons, particularly involving the congruence of semantic categories and the calquing of lexical items across languages. Calquing is especially pervasive in place names and ethnonyms, binomial names for flora and fauna, and items of material and ritual culture (e.g. the name of a regional culture hero: Tukano o’â-kô ‘Bone-Son’, Tariana yapî-rikû-ri ‘One on the Bone’, Hup g’eg teh ‘Bone Son’); see Aikhenvald (2002: 229), Epps (2013), Floyd (2013).

Grammatical structures and categories among the Vaupés languages have been profoundly affected by areal diffusion, giving rise to a significant degree of morpheme-to-morpheme and word-to-word intertranslatability (see Aikhenvald 2007a: 261). East Tukanoan languages have provided the model for many of these changes, but in some cases they too have converged to be more like their neighbors. The extent to which East Tukanoan languages have been shaped by contact with each other is not as easy to determine, but diffusion among these languages has also undoubtedly occurred (Gomez-Imbert 1993, Chacon 2013).

The heavy restructuring experienced by Tariana, the only Arawak language fully incorporated into the Vaupés linguistic exogamy system, has been explored in detail by Aikhenvald (1999a, 2001b, 2002, 2007a, inter alia). As Aikhenvald demonstrates through comparison with Baniwa, Tariana’s closest Arawak sister outside the Vaupés region, Tukanoan influence has led to changes throughout Tariana grammar, including a realignment of morphemes and constituent order; the development of case-marking on nominal arguments (topical non-subject and oblique, as well as the reduction of multiple locative markers to one catch-all marker); the elaboration and restructuring of the nominal classifier system; the augmentation of the evidential system from two categories to five; the development of verb compounding, switch-reference, new complementation strategies, discourse marking, etc. Tariana also exhibits significant convergence in its phonological inventory towards the typical Tukanoan profile, diverging from typical Arawak profiles (Chang and Michael 2014: 1-2). Examples (1-3) (from Aikhenvald 2007a: 245-246) illustrate the near-isomorphism between Tariana and Tukano, in contrast to Tariana’s sister Baniwa, as can be seen in the following examples.

(1) Tariana (Arawak)
\[\text{nese pa:ma di-na} \quad \text{du-yana-sita-pidana} \]
\[\text{then one+NUM.CL ANIMATE FEM 3sgf-OBJ ALREADY-REM.PAST.REP} \]
\[\text{‘She had reportedly cooked him already.’} \]

(2) Tukano (East Tukanoan)
\[\text{tiîta ni’kô kî-re} \]
The Nadahup languages have also undergone significant restructuring through Tukanoan influence, facilitated by the one-way bilingualism that predominates among many Nadahup speakers in the Vaupés. Hup appears to have been most profoundly affected (see Epps 2005, 2007a, 2008, *inter alia*), having developed prosodic nasalization, verb compounding, nominal classification, evidentials (as with Tariana, a five-way distinction built atop an earlier two-way one), future and recent/distant past tense distinctions, new strategies of number marking, and case-marking (including an object/non-subject marker that is sensitive to animacy and definiteness, and a catch-all locative), among many other features (compare example 4 to those above). Hup’s sister Yuhup, also spoken in the Vaupés, has likewise experienced significant Tukanoan influence (with diffusion-induced changes probably beginning in their common ancestor), while Dâw, located on the periphery of the Vaupés, has been less affected. Nadëb, their most distant sister, has been essentially unaffected by diffusion from Tukanoan (though it has almost certainly been influenced by Arawak), and its grammar is strikingly different from those of its sisters.

(4) Hup (Nadahup)

\[
yɨ \ tɨy=ʔåy \ tɨh-ân \ \text{chw}-yiʔ-\text{chw}îy=\text{mah} \ j'ẫm
\]

‘Then she had reportedly cooked him already.’

Kakua (Kakua-Nukak) likewise shows evidence of diffusion, probably also from Tukanoan languages. As can be seen in example (5) (from Bolaños 2012: 3), Kakua resembles its neighbors with respect to characteristics such as evidentiality, verb compounding, a recent/distant past tense distinction, non-subject case marking, and verb-final constituent order. Chang and Michael (2014) also find evidence of phonological convergence between Kakua and the Nadahup languages Hup and Yuhup.

(5) Kakua (Kakua-Nukak)

\[
\text{hiv} \ \text{kan-ðiʔ} \ \text{ʔå-t-hêm-ep-wît-be}
\]

‘The jaguar reportedly ate him.’

Arawak languages have also exerted influence on some East Tukanoan languages. Evidence for diffusion in this direction includes the development of aspirated stops in Kotiria (Wanano), possessive proclitics in Kotiria, Kubeo, Tatuyo, and certain
other languages, and the use of shape classifiers with non-human animates in Kubeo and (to a limited extent) in Kotiria (Stenzel and Gomez-Imbert 2009, Gomez-Imbert 1996).

A number of further studies have undertaken in-depth comparative investigations of particular grammatical features across Vaupés languages, noting their many similarities; these have explored serial verb constructions, including grammaticalization of the verbs *come/go* as directionals (Ospina and Gomez-Imbert 2013, Aikhenvald 1999b); the expression of spatial relations (Stenzel 2013b, Epps and Neely 2014); possessive constructions, including the development of a comparable alienable/inalienable distinction and similar marking strategies (Stenzel 2013a); differential object marking strategies (Zuñiga 2007, Stenzel 2008); and nominal classification (Aikhenvald 2007b, Epps 2007b).

The Vaupés languages provide intriguing insights into the mechanisms by which areal diffusion takes place. In many instances, these mechanisms have involved the grammaticalization of native material to create new categories and structures. The process can involve the calquing of whole constructions, as can be seen for example in the close parallels in event packaging via serial verb constructions. Rather than resulting in simplification, the development of new material through contact-driven grammaticalization often produces an overall increase in grammatical complexity by adding to existing repertoires of categories and structures, as in the case of Tariana’s development of case-marking in addition to its earlier strategy of pronominal cross-referencing (see Aikhenvald 2003). Stenzel’s (2013a) discussion of possession-marking strategies among Vaupés languages highlights the fact that even while distinct languages may converge on a common model through processes of diffusion, the language-internal dynamics of change give rise to fine-grained differences, such that the outcomes are not isomorphic. Epps (2012) shows that despite the constraints against the borrowing of form in this region, perceived formal similarities nevertheless can and do play a role in facilitating structural adaptations, as in the case of the development of an evidential form *ni* from an existential verb in Tukanoan and Hup, and in Tariana from a functionally distinct marker * nhi* ‘anterior tense’ (Aikhenvald 2002: 123).

Particular examples of diffusion-driven change in Vaupés languages also highlight the role of discourse in motivating these processes (compare also Aikhenvald’s 2007a: 261 observation that “the more pragmatically motivated, the more diffusible”). The development of evidentials provides a good example: norms of conversation, story-telling, and other discourse forms, shared among speakers, foster an expectation that one’s source of information will be explicitly stated (and that if it is not, one’s reliability or responsibility may be in question). In Hup, this discursive expectation apparently led to an increase in the frequency of verb roots associated with information source appearing in verbal compounds encoding events, which in time grammaticalized into evidentials (see Epps 2005). Moreover, regional norms for narrating traditional stories may have given rise to a further change in Hup: In these stories, almost every clause is marked by the reported evidential *=mah*, followed directly by the distant past tense marker *j’am* (itself a fairly recent addition to Hup grammar), as can be seen in (6) – although in other discourse contexts the past markers only appear when the temporal information is contrastive or particularly emphasized. Moreover, in one dialect of Hup, the reported evidential appears fused together with the distant past marker in traditional
stories to produce =máam (example 7). Given that Hup has almost no other portmanteau morphs, the parallel with the fused tense-evidential suffixes in Tukanoan languages is striking.

(6) Hup, Middle Tiquié River

\[ γίνη = māh \quad j'ām \quad tīh \quad bīʔ-ʔ, \]
so=REPORTED DIST.PAST 3sg make-DECL

\[ hūp = n'ān \quad tīh \quad bīʔ-ʔh \]
person=PL.OBJ 3sg make-DECL

‘Thus (long ago, they say) he made (them), he made people.’

(7) Hup, Lower Tiquié River

\[ j'ūg-ūt = maám \quad tīh \quad wɔn-kot = māh-ah \]
forest-OBL=REP.DIST.PAST 3sg follow-go.in.circles=REP-DECL

‘In the forest (long ago, they say), he wandered following (the tapir).’

2.2 The Caquetá-Putumayo region

The Caquetá-Putumayo area of southern Colombia and northern Peru is another local diffusion zone, though it is not nearly as well studied as its northern neighbor the Vaupés. In this region, speakers of languages belonging to the Bora and Witoto families, the Arawak language Resigaro, and the isolate Andoke interact intensively through marriage, ritual contexts, etc.; refer to themselves together as the ‘People of the Center’; and are widely multilingual4 (see Echeverri 1997, Seifart et al. 2009, Londoño Sulkin 2012). Distinctive shared cultural practices include the ritual ingestion of powdered coca leaves and liquid tobacco, extensive song cycles, and particular styles of warfare, personal hygiene, etc.; at the same time, the groups have maintained distinct languages, origin stories, and certain other emblems of identity (Seifart 2011: 7-8, Whiffen 1915).

As in the Vaupés, the People of the Center share an “inhibition against lexical borrowing” (Seifart 2011: 88). Despite their close contact, Seifart (2011: 20) finds that only about 5% of Resigaro lexical stems have been borrowed from Bora (gauged via an extensive set of core and non-core vocabulary), of which many are flora-fauna terms; similarly low loan rates between Bora, Resigaro, and other languages of the region are identified in work by Bowern et al. (2011, 2014).5 However, Resigaro has borrowed a striking number of bound morphological forms, including whole sets of nominal classifiers (over half of Resigaro’s total), number markers, quantifiers, and other forms, totaling over 50 distinct items (Seifart 2011, 2012). Further diffusion has affected grammatical structures and categories in Resigaro, with or without the mediation of directly borrowed morphemes.

4 However, as in many of the other regions discussed in this paper, recent decades have seen significant language loss as speakers shift to Spanish (or other colonially mediated languages).

5 The figure of 24% loans in Resigaro given by Aikhenvald (2001b: 182; see also Eriksen and Danielsen 2014: 188) was probably erroneously inflated by the inclusion of borrowed classifier forms.
features include the development of an inclusive/exclusive distinction, second-position tense-aspect-mood clitics, the loss of object cross-referencing suffixes, and the restructuring of verbal morphology (Seifart 2011: 14, Aikhenvald 2001: 189). Resigaro phonology has also been affected, with the addition of new phonemes (/ɸ/, /dʒ/, /ʔ/), syllable structure restrictions, and a two-tone contrast (Aikhenvald 2001b, Seifart 2012, Chang and Michael 2014). Indeed, Chang and Michael (2014) show that Boran and Witotoan languages, Resigaro, and the isolate Andoke exhibit significant convergence in their phonological inventories, allowing us to pick out the People of the Center as a well-defined phonological area.

The People of the Center have been involved in longer-range processes of diffusion as well. They are in contact with the Arawak Yucuna to the north, which themselves are in close contact with the East Tukanoan Tanimuca (Retuarâ; see Seifart 2007, Aikhenvald 2001b); the Caquetá-Putumayo groups share a number of characteristics with their northern neighbors in the Vaupés. These similarities are both cultural (e.g. large signal drums) and linguistic (e.g. the distributional and functional properties of nominal classifiers, nominative-accusative alignment, etc.; see Aikhenvald 2001b: 189, Seifart and Payne 2007).

Figure 1 models a subset of northwest Amazonian languages as a NeighborNet splitsgraph, with respect to 226 grammatical features (mostly
The splitsgraph illustrates the extent to which areal diffusion has produced regional grammatical profiles among the Vaupés and Caquetá-Putumayo languages. The Vaupés cluster includes East Tukanoan, Hup, and Kakua, while the Arawak languages Tariana, Resigaro, and Yucuna pattern more closely with their non-Arawak neighbors than they do with their closest northern Arawak relatives. Also included are Kokama and Nheengatú (both Tupí-Guaraní) and Yagua (Peba-Yaguan); the association between these three languages is undoubtedly indicative of areal diffusion in the region where Yagua and Kokama are spoken, just south of the Caquetá-Putumayo area.

---

6 See www.laits.utexas.edu/huntergatherer.

7 We note that the tendency for grammars to change through areal diffusion while lexicons remain relatively conservative, as seen in these and other Amazonian languages, presents a challenge for the view that typological features will in general be more likely to retain a deep-time genetic signature (e.g. Dunn et al. 2005, Sicoli and Holman 2014).
2.3 The Upper Xingu Region

Contact among the Upper Xingu peoples has been relatively well documented ethnographically, but its linguistic effects have only begun to be explored. This region is home to more than a dozen languages belonging to the Carib, Arawak, Jê, and Tupi families, as well as the isolate Trumai, although several of these groups are relative newcomers (arriving between the 16th and 20th centuries in response to colonial pressures). Like the Vaupés and other Amazonian regional systems, the Xingu is an area of intense interaction among groups, particularly through ritual contexts and trade. The groups share a distinctive “cultural package” (Fausto et al. 2008: 137; see also Franchetto 2011), with particular rituals, hairstyles, house architecture, etc., to which newcomer groups have assimilated as they were absorbed into the Xingu system (see e.g. Guirardello-Damien’s 2011 discussion of the Trumai).

As in the Vaupés, language plays an important role in the Xingu as a “basic diacritic” of ethnic identity and place within the regional system (Fausto et al. 2008: 141). Speakers provide explicit characterizations of linguistic differences,
observing for example that the Carib groups speak ‘in the throat’ or ‘inwards’, while the Arawak peoples of the region speak ‘outwards’, ‘on the tip of the teeth’ (Fausto et al. 2008:143). These “rigorous and active processes of differentiation” are realized via a monolingual everyday ethos (Ball 2011: 93; see also Seki 2011: 85), in contrast to the Vaupés; yet multilingualism is extensive in ritual contexts, and the frequency of interaction has led to characterizations of the region as a “communications network” (Basso 1973: 5) or even a speech community (Ball 2011: 93). Little code-switching and relatively little lexical borrowing occur (though some loans exist; see Seki 1999, 2011).

Seki (1999, 2012) has characterized the Xingu as an “incipient” linguistic area, primarily on the basis of lexical similarities and common features of myth and ceremonial discourse. She identifies several grammatical features that have diffused within the region, and it is likely that further linguistic exploration will reveal more; these include the loss of a masculine-feminine gender distinction in the Arawak languages of the region (in cross-referencing and in independent pronouns), the development of a phoneme /ɨ/ in Arawak, a p > h shift in Carib and Tupi-Guarani, and a change to CV syllable structure in Carib. Chang and Michael (2014) clarify the multilateral nature of phonological borrowing in the Xingú area, confirming the diffusion of /ɨ/ into Xingu Arawak languages from their Carib or Tupi-Guarani neighbors, and adding the diffusion of /ts/ into the Xinguan Carib languages from their Arawak neighbors and the diffusion of nasal vowels into the Xinguan Arawak and Carib languages from their Tupi-Guarani neighbors. Lexical restructuring and calquing have also taken place, including the development of comparable systems of post-nominal elements meaning ‘big, supernatural, hyper’, ‘similar to’, ‘true/genuine’, and ‘bad/worthless/unsatisfying’, which are used productively to create new complex nouns. For example, in Yawalapiti úi ‘snake’ + kumã ‘hyper’ yields úi-tyumã [kumã] ‘snake-spirit’ (Viveiros de Castro 2002), and in Trumai fi ‘cigar, cigarette’ (tobacco)’ + yuraw ‘hyper’ (itself a loan from Tupi-Guaraní Kamayurá) yields fi yuraw ‘marijuana cigarette’ (considered abnormal, dangerous; see Guirardello-Damien 2011: 120). The diffusion among the Xingu languages appears to be generally multidirectional.

2.4 The Guapore-Mamore region

The Guaporé-Mamoré of southwest Brazil and northeast Bolivia is home to over fifty languages from a wide range of families (Arawak, Macro-Jê, Chapacuran, Tupí, Nambikwara, Pano, Tacanan, and many isolates). Ethnographic documentation indicates that many of these groups have engaged in extensive interethnic contact, intermarriage, and exchange, yielding a regional culture area with commonalities in territorial subgroups, bodily adornment, mythological themes, etc. (Levi-Strauss 1948, Maldi 1991).

The investigation of contact among the Guaporé-Mamoré languages has not yet been extensive, but work by Crevels and van der Voort (2008; see also van der Voort 2005) indicates that diffusion has certainly taken place and that this is a rich area to explore. As in the other regions considered here, levels of lexical borrowing among these languages appear to be quite low (estimated at around 5% by van der Voort 2005: 395; see also Crevels and van der Voort 2008: 164); however, they share many structural similarities, including evidentials, an inclusive/exclusive
distinction, a high incidence of prefixes, and a strong preference for verbal number (i.e. alteration of the verb via suppletion, reduplication, or affixation to express the number of the subject or object), accompanied by a general lack of nominal number (Crevels and van der Voort 2008: 167). Further similarities include verbal cross-reference systems with similar morpheme positions, and rich systems of directional morphemes, such as Kwaza’s ‘movement in a circle’, ‘into fire’, ‘behind the house’ (van der Voort 2005: 399; see also Guillaume forthcoming). Eriksen and Danielsen (2014: 175) also note areal influence in verbal morphology and personal pronominal distinctions in two Arawak languages of the region (Paunaka and Moxo).

A particularly interesting aspect of areal diffusion in the Guaporé-Mamoré involves the direct borrowing of morphological forms, particularly within nominal classifier systems (van der Voort 2005: 397, Crevels and van der Voort 2008: 167) – a striking parallel to the effects of contact in the Caquetá-Putumayo. Several of these forms are given in Figure 2; none of the five languages listed are known to be related.

<table>
<thead>
<tr>
<th></th>
<th>‘bark’</th>
<th>‘fruit’</th>
<th>‘bone’</th>
<th>‘tooth’</th>
<th>‘liquid’</th>
<th>‘round’</th>
<th>‘thorn’</th>
<th>‘porridge’</th>
<th>‘powder’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kwaza (isolate)</td>
<td>-kalo</td>
<td>-ko</td>
<td>-s</td>
<td>-mũ</td>
<td>-mũ</td>
<td>-tɛ</td>
<td>-nĩ</td>
<td>-mɛ̃</td>
<td>-nũ</td>
</tr>
<tr>
<td>Kanoê (isolate)</td>
<td>-ko</td>
<td>-s</td>
<td>-mũ</td>
<td>-tɛ</td>
<td></td>
<td></td>
<td></td>
<td>-nũ</td>
<td></td>
</tr>
<tr>
<td>Aikana (isolate)</td>
<td>-zu</td>
<td>-mũj</td>
<td>-mũ</td>
<td>-haw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arikipú (Macro-Jê)</td>
<td></td>
<td></td>
<td>-nĩ</td>
<td>-mrɛ</td>
<td>-nũ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nambikwara (Namb.)</td>
<td>-kalo</td>
<td>-su’</td>
<td></td>
<td></td>
<td>-nũx3</td>
<td></td>
<td></td>
<td>-nũx3</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Similar classifier forms in Guaporé-Mamoré languages (van der Voort 2005: 397)

As Muysken et al. (forthcoming) point out, the fact that so many of the Guaporé-Mamoré languages are small families or isolates located only within this region makes it difficult to test for contact by comparison with relatives outside. However, these authors apply quantitative measures to demonstrate that at least some features must have converged via diffusion in this region. Other features that are more widely attested may themselves be the outcome of more far-reaching contact networks that existed in the past, as discussed in Section 3 below.

2.5 Other regional diffusion zones in lowland South America

Several other localized diffusion zones exist in the Amazon basin and adjacent lowland regions, and still others probably remain undetected. One area that has had some documentation is the Southern Guiana region, where Carib, Arawak, and Salivan groups exhibit notable cultural continuity resulting from “constant interaction through marriage, trade, and migration” (Rivière 1984: 8; see also Arhem 1989, Carlin 2011: 226). Much like the other regions discussed here, linguistic distinctions are maintained as key markers of ethnic identity (Howard 2001: 341), and loanwords are relatively few (Carlin 2007, Bowern et al. 2011, 2014). Carlin (2007) discusses grammatical diffusion from Carib languages
(principally Trio and Waiwai) into the Arawak language Mawayana, established in part via comparison with Mawayana’s sister Wapishana. Contact-induced changes in Mawayana include the addition of a first person exclusive distinction (via borrowing of a Waiwai pronominal form), the development of nominal tense marking, affective marking on nouns or verbs to express ‘pity’ or ‘recognition of unfortunate circumstance’, a frustrative marker on verbs (indicating that the action was carried out ‘in vain’), and a ‘similative’ or ‘as if’ marker on nominals (see examples 8-9, from Carlin 2007: 329). Carlin notes that all of these categories are obligatory in the local Carib languages.

(8) Mawayana (Arawak)

waata-ni r-ayādyā
opossum-simil 3s-transform.past
‘He changed into an opossum.’

(9) Trio (Carib)

kaikui-me iēmeta
jaguar-simil he.transformed
‘He transformed into a jaguar.’

The Southern Guiana region may itself be part of a much larger contact zone that includes parts of the Orinoco and northern Amazon watersheds (Migliazza 1985: 20; cf. Campbell 2012: 306). Further possible diffusion zones include the area of the Tocantins and Mearim Rivers in northeastern Brazil, home to Jê and Tupí-Guaraní languages (Braga et al. 2011), and the Venezuelan-Antillean area (Constenla Umaña 1991: 125-126, Campbell 2012: 307).

Finally, the region of the Gran Chaco, located just to the south of the Amazon basin, bears striking similarities to the Amazonian regional ‘systems’ discussed here: The speakers of its six distinct language families share many cultural similarities (Braunstein and Miller 1999: 9-11), shun code-switching and lexical borrowing (Campbell and Grondona 2010, 2012: 657; Vidal and Nercesian 2009: 1023), practice a form of linguistic exogamy, and identify strongly with a single language while understanding many (Campbell and Grondona 2010). The Chaco languages share many structural features, including animal classifier(s), active-inactive verb alignment, similar strategies of pronominal affixation, and rich demonstrative systems, many of which are undoubtedly due to diffusion (Campbell and Grondona 2012).

We will probably never know the extent to which the dynamics of these regional diffusion zones applied more generally in the Amazon basin, across both space and time. However, the wide distribution of regional ‘systems’ and their notable social and linguistic parallels suggest that these are not isolated cases. If Eriksen and Danielsen (2014: 163) are correct in their assertion of “a vast socio-religious and economic exchange system that affected the lives of all inhabitants of northern South America between 1000 BCE and 1000 CE,” then longer-range diffusion would undoubtedly have linked these local zones within wider networks. We turn to this question below.
3 Long-range language contact and macro-areality in Amazonia

The linguistic areas discussed in Section 2 generally arose through interactions between neighboring peoples in networks that span relatively small geographical regions in comparison to Amazonia as a whole. There is clear evidence, however, that Pre-Columbian Amazonian peoples were linked via large-scale social, ritual, and trade networks that spanned considerable areas of the continent, raising the possibility that Amazonia may exhibit linguistic areality on a similarly large scale. The fact that many cultural features and practices, such as perspectivism in mythology (Viveiros do Castro 1998, 2012: 48-51), the use of sacred flutes (Hornborg and Hill 2011: 17), and a variety of subsistence techniques (see e.g. Carneiro 2000 on manioc-processing technology), have diffused over large areas of Amazonia suggests that linguistic features may have diffused on a similar scale.

Although the full extent of pre-Columbian South American trade networks is impossible to recover at this point, ethnohistorical and archeological evidence clearly indicates that such networks spanned large areas of Amazonia, and linked Amazonia to adjacent regions, such as the Andes and Chaco. The earliest accounts of the colonial period provide ample evidence of vast trade networks criss-crossing the content. Citing Oviedo y Valdés (1851-1855[1535]), for example, Nordenskiöld (1922: 7) describes Arawak traders then living near the mouth of the Amazon making 1500 kilometer trading journeys along the coast in groups of 50-60 canoes and 500-800 men. He similarly discusses evidence of trade between the Cariban peoples of the Guianas and the Cariban Carijona of the Putumayo, and further still to peoples on the Amazon proper (ibid: 149-150). Trade links between the Amazon and the Guianas are echoed by Fritz (Edmunson 1922) who indicates that the Omaguas, who occupied much of the Amazon between the Japurá and Napo, formed a node in a trade network linked to the Guianas, and traded with peoples deeper into the Amazonian headwaters regions. Nordenskiöld presents evidence of similarly long-distance trade between the Guaranian peoples of the Paraná and Paraguay River Basins and the Andean Inca Empire (ibid: 7-10, 133-134). Trade of specific rare products is similarly known to have extended over thousands of kilometers, as in the case of salt – mined by the Arawak Ashéninkas in the Peréné River basin in the Andean foothills, and compellingly argued by Rydén (1962: 652) to have reached the Tupinambas who had resettled on the Amazon at Tupinambarana, downriver of the mouth of the Río Negro. Eriksen and Danielsen (2014) cite ethnohistorical and archeological evidence to argue for the existence of an extensive Arawak-dominated trade network spanning much of western Amazonia, extending the arguments of Vidal (2000), who provides evidence for a major Arawak-dominated pre-Columbian trade network in northwestern Amazonia.

Long-distance trade of this sort survived well into the modern period. Lathrap (1973), for example, describes the varied networks in which the Shipibo communities, located on the Ucayali River, participated. In some of these networks, raw materials for ceramic production circulated up to 240 kms from the communities studied. Other networks were even more expansive and interethnic, and included Yagua communities some thousand kilometres downriver, through which Shipibos obtained blowguns, as well as Tikuna communities, located a further five hundred kilometres along the Amazon proper, from which they obtained blowgun darts. Roth (1924) similarly describes a network spanning the
eastern Orinoco basin and the Guianas, characterized by circulation of rare materials and high degrees of craft specialization (see also Dumont 1991). This network linked the Tupí-Guaraní communities on the Oyapock, producers of valued grinding stones; the river-craft-producing Warao at the mouth of the Orinoco; the Cariban Waiwai, located along the hills separating the Guianas from the Amazon watershed, and who specialized in a number of palm products; and the Cariban Makiritare, located in the central Orinoco basin, who produced hammocks, cassava graters, and ornamental products.

Whether far-flung networks of interaction like these have led to widespread grammatical borrowing and convergence, and to the emergence of linguistic macro-areas within Amazonia, has become a topic of increasing prominence in Amazonian linguistics, as has the question of whether Amazonia as a whole constitutes a linguistic area. One of the earliest macro-area proposals considered the region encompassing the Orinoco River Basin and the portion of the Amazon Basin containing the northern tributaries of the Rio Negro (Migliazza 1985). Features that Migliazza attributed to this area included ergative alignment, OV order, lack of a passive construction, relative clauses formed by apposition and nominalization. Most Amazonianists would now recognize this list as including several features of broader Amazonian distribution, and indeed, the features that Derbyshire (1987: 311) tentatively proposed as defining an Amazonian linguistic area include most of these, and in addition, a preference for OS order, subject and object verb agreement, null free argument realization, head-modifier order, and complex morphology. Derbyshire and Payne (1990) subsequently added noun classifiers to this list of tentative features.

Perhaps the most promising macro-area proposals have been based on typological divisions bisecting the east-west axis of the continent. In one of the earliest such proposals, Doris Payne (1990: 5) suggested the existence of a Western Amazonian area consisting of the lowland areas to the west of the Andes characterized phonologically by complex stress and ‘pitch accent’ systems and morphosyntactically by a strong tendency towards polysynthesis and complex verbal morphology, directional, locational, and positional morphemes, and a distinctive type of noun classification system. The validity of this proposal is somewhat difficult to evaluate because neither the precise limits of the area, nor the distribution of indicated features inside and outside the proposed area are given; but as discussed below, recent quantitatively based work lends support to an east-west areal split in South America.

As descriptions of Amazonian languages increase in both number and quality, work identifying macro-areas on the basis of relatively fine-grained linguistic phenomena will probably become more common. A promising example of this type is Guillaume and Rose’s (2010) suggestion that sociative causatives may be an areal feature of southwest Amazonia, with the distribution of sociative causatives outside this area attributed to the spread of Tupian languages from their southwestern homeland. A systematic examination of the distribution of such morphemes both within South America and beyond is an obvious target for future research.

The question of linguistic macro-areas within Amazonia naturally leads to an issue already raised in our discussion: whether Amazonia as a whole constitutes a linguistic area. Although this remains an open question, an emerging consensus
points to Amazonia not forming a linguistic area *sensu strictu*. Dixon and Aikhenvald (1999: 7-10) is perhaps one of the best known and most explicit efforts to enumerate characteristics that define Amazonia as a linguistic area, and as complementary to an Andean linguistic area. The grammatical features proposed to be ‘shared by all (or most)’ Amazonian languages range from polysynthesis and head-marking, to TAM categories being expressed as optional suffixes, to adverbs and adpositions being incorporable into verbs. As Birchall (2014) observes, however, the empirical basis for the claimed areality of these features is unclear (as, indeed, they are for Derbyshire's (1987), Migliazza's (1985), and Payne's (1990) proposals, discussed above), raising the need to move beyond impressionistic claims regarding areality to explicit quantitatively-grounded methods that make use of suitably organized, sufficiently large and dense, and ideally, publicly available datasets (Haspelmath 2004).

Steps towards databases of these types for South America include the South American Indigenous Language Structures database (Muysken et al. 2014), and the South American Phonological Inventory Database (SAPhon; Michael et al. 2012), and quantitatively sophisticated work based on these resources has recently begun to appear. Significantly, these works support not an Andean–Amazonian areality split, that leaves Amazonia as a clearly defined area, but a different west-east split – where the western area corresponds roughly to the Andes, Southern Cone, and Doris Payne's Western Amazonian region, while the other large linguistic area to the east consists of the remainder of the continent (see also van Gijn et al., this volume).

One illustration of this east-west division can be seen in Birchall's (2014) study of argument marking features in 74 South American languages, in which he tests for a statistically significant concentration in each of seven South American geographical macro-areas. Echoing Krasnoukhova’s (2012) qualitative results on noun-phrase features in South America, Birchall’s analysis suggests that the features he examined pattern similarly within an Eastern South American Linguistic Area (ESALA) on one hand, comprising northern and southern Amazonia and the Chaco-Planalto area, and within a Western South American Linguistic Area (WSALA) on the other, comprising the north and central Andes, western Amazonia, and the Southern Cone. For example, Birchall (ibid: 203) finds that ergative alignment, suggested by Dixon and Aikhenvald (1999) to be a general Amazonian feature, has a statistically significant association only with the southern Amazon region and ESALA, and not with northern or western Amazonia. Similarly, clusivity distinctions, proposed by Adelaar (2008: 29) to be an Andean feature, do not emerge as particularly Andean, but once again, as WSALA feature (Birchall 2014: 205-206). In contrast, several features, such as the use of both indexation and case as argument marking strategies, and accusative alignment for NP and pronoun arguments, turn out to be not mainly Andean, as proposed by Dixon and Aikhenvald (1999), but WSALA characteristics, providing further evidence for an E/WSALA areality split rather than an Andean/Amazonian one.

These results are broadly congruent with quantitative computational work on phonological areality in South America, which finds that lowland languages exhibiting phonological similarities to Andean ones – for example, uvular and ejective consonants, palatal laterals, multiple liquids, and small vowel inventories – cluster near the central Andes and in the Southern Cone (Michael et al. 2014),
forming a phonological area corresponding roughly to WSALA. Similarly, languages in the remainder of the continent, corresponding approximately to ESALA, exhibit larger vowel inventories that include mid and high central vowels, nasal vowels or supersegmental nasality, labial fricatives, and a glottal stop, among other features.

Despite the lack of support for a general Amazonian linguistic area evident in these results, there are nevertheless phenomena that are found in many Amazonian languages, although these are not pervasive enough to be diagnostic of a linguistic area in the usual sense of the term. In recognition of this fact, Aikhenvald (2012: 68-71) has more recently introduced the term ‘language region’, in contrast to the more rigorously defined notion of ‘linguistic area’, to characterize Amazonia as a whole. The precise explanation for recurrent but sporadically attested ‘Amazonian’ features, such as antipassives and complex classifier systems (ibid.: 70), is unclear, but there is suggestive evidence that in some cases, linguistic items and features have diffused along long-range networks of the kind discussed earlier in this section. For example, Epps (2014; see also Bowern et al. 2014) identifies several dozen Wanderwörter that indicate the widespread diffusion of terms associated with important animal and plant species and food items across much of northern Amazonia. Epps also observes a widespread tendency across Amazonia for numeral terms indicating ‘4’ (and occasionally ‘3’ and ‘5’) to be formed using ‘relational’ nouns or verbs, usually meaning ‘companion’ or ‘accompany’, a strategy that is extremely rare outside of Amazonia (Epps et al. 2012, Epps 2013). Epps suggests Tupí-Guaraní languages in particular may have played an important role in the diffusion, but observes that the ultimate source of these apparently diffused items remains an open question. Regardless, the fact that items like these have circulated widely but sporadically across Amazonia suggests that borrowing may be mediated by sparse networks that link relatively distant languages without directly affecting closer neighbors.

Another line of research addressing Amazonian areality seeks to identify areal patterns not in grammatical structure, but in discourse practices and language ideologies. Beier et al. (2002), for example, argue that particular discourse practices, such as the dialogical discourse genres, ritual wailing, and the pragmatically-motivated use of evidentials are common over large areas of Amazonia in a manner consonant with that of a language region. Bowern et al. (2011) similarly find that lexical borrowing is unusually low in Amazonia, in comparison with other global macro-regions (see also Section 2 above), a tendency that Epps (forthcoming) attributes to a widely diffused language ideology that discourages language mixing.

4 Conclusion

Amazonia offers important insights into the dynamics of language contact and the development of areal linguistic patterns. As this chapter has explored, striking similarities exist among regional contact zones throughout the Amazonian lowlands, in which intense interaction and a degree of cultural homogeneity tend to go hand in hand with high linguistic diversity. In these regions, language is afforded special salience as a marker of distinct social identities, and thus while
individuals are often highly multilingual, there is a marked absence of phenomena that are commonly associated with multilingualism in other parts of the world – most notably code-switching and lexical borrowing. Similarly, linguae francae and clear examples of language shift (aside from those associated with the expansion of the European and Quechuan spheres), while common in many other regions of the world where multiple languages are represented, appear to be rare or even unattested in these Amazonian contact zones. At the same time, the diffusion of grammatical categories and structures across languages is ubiquitous in these regions, although for most of these contact zones these processes have only begun to be explored. In some cases, the grammatical borrowing even includes the transfer of bound morphology, most notably that associated with nominal classification. This prevalence of extensive grammatical diffusion coupled with restrained lexical borrowing is typologically significant, since current conceptions of language contact dynamics stress the importance of lexical mediation in the diffusion of grammatical material, and particularly of bound morphology.

On a larger scale, Amazonian linguistic areality provides some intriguing glimpses into the dynamics of human interaction in prehistory, when extensive trade routes would have criss-crossed the Amazon basin and linked it with other parts of the continent. The broad east-west division outlined above suggests that social networks linking the Andes, the western lowlands, and the Southern Cone on one hand, and the eastern and central lowlands on the other, would have been particularly active, whereas networks that set the Amazon basin apart from other regions may have been less significant or functioned in different ways.

The study of areality in Amazonia and South America more generally has entered an exciting new phase. The empirical basis for such studies has been greatly enriched by the recent surge in basic descriptive work, which in turn is feeding a number of databases that will facilitate systematic study. Likewise, new analytical and theoretical approaches to areality are being developed through the application of quantitative techniques and via holistic approaches that take geography, demographics, and culture into account. We look forward to exciting findings in the coming decades, and to the new insights these will provide into South American prehistory and the mechanisms and processes involved in language contact.

References


Gomez-Imbert, Elsa. 1996. When animals become ‘rounded’ and ‘feminine’: Conceptual categories and linguistic classification in a multilingual setting. In


Seifart, Frank, Doris Fagua, Jürg Gasché, and Juan Alvaro Echeverri (eds.). 2009. A multimedia documentation of the languages of the People of the Center. Online publication of transcribed and translated Bora, Ocaina, Nonuya, Resígaro, and Witoto audio and video recordings with linguistic and ethnographic annotations and descriptions. Nijmegen: DOBES-MPI.


