

What Is Riau Indonesian?

David Gil

Max Planck Institute for Evolutionary Anthropology, Leipzig

ABSTRACT

This paper addresses the question: What is Riau Indonesian? Previous studies of Riau Indonesian have attracted a variety of criticisms suggesting that, in some sense or another, it is not a "real" or "proper" language. This paper takes up and dismisses 12 specific claims regarding the nature of Riau Indonesian, including, among others, that it is a corrupt, broken, imperfect language variety, a language variety without native speakers, an artefact of code switching, or a creole. Examination of Riau Indonesian in sociolinguistic space, in relationship to its substrate and superstrate languages, and in geographical space, in relationship to neighboring varieties of colloquial Indonesian, suggests that there is nothing particularly exceptional about it, and that it is a run-of-the-mill language just like numerous others. In conclusion, it is suggested that other major world languages may also possess a range of language varieties, similar in their broad sociolinguistic profiles to Riau Indonesian, but which in some cases may not yet have been adequately recognized or described.

1. IN SEARCH OF A NAME, IN SEARCH OF AN IDENTITY

Languages and dialects do not present themselves to us with ready-made names, well-established identities, and their own individual profiles plus three-letter codes in the latest edition of *Ethnologue*. Often, several distinct languages or dialects share a single name; conversely, a single language or dialect may be known by several different names, or, alternatively, may not have any name of its own. Linguists try to clear up the mess, by engaging in careful descriptions, both of linguistic structure (lexicon, phonology, morphology, syntax, and so forth) and of the sociological and geographical landscape in which such structure is embedded. However, in order to describe a language or dialect, the linguist must have some prior notion of what that language or dialect is, some presupposed delimitation of the object of inquiry. The identification of languages and dialects thus involves a continual back-and-forth interplay, with delimitation informing description which in turn contributes to further delimitation, and so on.

My first serious encounter with these issues was in 1992, when I took up a position at the National University of Singapore. My office window opened out onto a vista of beckoning tropical islands, part of the Riau archipelago of neighboring Indonesia, and I soon found myself visiting these islands on a regular basis, and picking up the local language. But what language was that? Indonesian, of course; that is what people said that they were speaking, and it certainly was some form or another of colloquial Indonesian. However, after a very short time it became evident that this was a very different variety, in just about every respect — lexically, phonologically and grammatically — from the Indonesian described in textbooks and familiar in its broad outlines to many general linguists. Obviously, this was some kind of local basilectal language variety; but what exactly? Back in Singapore, I tried asking my local linguist colleagues about what it was they were speaking over in Riau, but nobody seemed to know or even care.

A few persons suggested that what I was encountering on my trips to Riau might be the renowned Riau Malay, that local dialect which, according to historians, formed the basis for the creation of the two standardized versions of the language: Standard Malay and Standard Indonesian. But this was clearly not the case. Indeed, in a handful of rather remote villages, the majority of the population were indigenous ethnic Malays, and what they spoke was in fact Riau Malay. However, a large majority of the inhabitants of the islands were actually migrants from other parts of Indonesia — Minangkabau, Batak, Javanese and others — and what they were speaking was clearly distinct from Riau Malay, lexically, phonologically and grammatically. Rather, it appeared to be some kind of lingua franca, or contact variety of Indonesian. Although seemingly dialects of the same language, the differences between Riau Malay and the local Indonesian were of sufficient saliency that speakers considered the two to be completely separate languages. In particular, whereas the local Malay was considered to be a distinct "regional language" (a concept for which Indonesian has a special term, *bahasa daerah*), on a par with other regional languages such as Minangkabau, Batak and Javanese, the local Indonesian was viewed to be a perhaps somewhat corrupted version of the "national language".

So this strange new variety of Indonesian was not Riau Malay, but it was just as obviously not Standard Indonesian either. Soon I was about to publish a first description of this language variety (Gil 1994), but it still had no name. What was I going to call it? Speakers simply referred to it as Indonesian, but that was not good enough. So I decided to call it "Riau Indonesian".

In the years since, I have published additional articles on Riau Indonesian, describing various aspects of its structure, and making various claims concerning its relevance for linguistic theory (Gil 1999, 2000, 2001a,b, 2002a,b, 2003b, 2004a,b, 2005c,d,e, 2006a, to appear). At the same time, I have gained a better understanding into its sociolinguistic character. In a nutshell: Riau Indonesian is the variety of colloquial Indonesian used in informal every-day contexts as a lingua franca for interethnic and increasingly also intraethnic communication by residents of the eponymous region. As a basilectal speech variety, it lies at the bottom of a lectal cline, or continuum, extending all the way up to the acrolectal Standard Indonesian. Riau Indonesian is distinct from, albeit in close contact with, other varieties of Malay/Indonesian spoken in the same region: Riau Malay, mentioned above, Bazaar Malay, a local variant of which is used for communication between speakers of Chinese and non-Chinese ethnicity, and Jakarta Indonesian, which is rapidly spreading through the country and acquiring the status of a pan-national mesolectal and somewhat "trendy" koiné. In its broad sociolinguistic profile, Riau Indonesian thus resembles many other regionally-based varieties of Malay/Indonesian used as basilectal lingua franche throughout the archipelago, such as Ambon Malay (van Minde 1997). One difference, however, is worthy of mention: whereas in eastern contact varieties such as Ambon Malay, the indigenous languages are at best distantly related to Malay/Indonesian, in many western contact varieties, including Riau Indonesian, the indigenous language is itself a dialect of Malay.

My work on Riau Indonesian has attracted a wide range of reactions. On the one hand, it featured in a one-hour documentary movie ("The Ways of Babel", by Pierre Morize, Arte France, Movimento Production, 2002), and was reported on in the *Economist* ("Babel's Children", 10-16 January 2004, pp. 61-62). On the other hand, it has triggered a number of adverse and sometimes quite outspoken reactions, mostly verbal rather than written. Some criticisms have been aimed at specific analyses proposed; however, in other cases, the negative reactions have questioned the validity of the data on which such analyses are based. Indeed, many of these reactions have revolved around the suggestion that, in one sense or another, Riau Indonesian is "not a real language".

The goal of this paper is to provide a refutation of such suggestions, by arguing that Riau Indonesian is indeed a real language, or at least as real as any other language. Section 2 of this paper, of an inevitably rather argumentative and polemic tone, formulates rebuttals of various specific claims concerning the nature of Riau Indonesian as something other than a proper language. Sections 3 and 4, of a more constructive and data-oriented flavour, examine the respective place of Riau Indonesian in sociological space, focusing on its relationship to its substrate and superstrate languages, and in geographical space, dealing with its relationship to its neighboring dialects. Emerging from the discussion is an answer to the question posed in the title of this paper: What Is Riau Indonesian? The answer is a rather mundane one: Nothing at all exceptional, just another one of the world's language varieties.

2. A DOZEN CLAIMS ABOUT RIAU INDONESIAN

What does it mean to say that Riau Indonesian is not a real language? For orthodox generativists, no language is real, since language, (or "E-language") is epiphenomenal, and what linguists should instead be concerned with is grammar (or "I-language"). But the people who say that Riau Indonesian is not real are not (just) generativists; they hail from all walks of linguistic life, ranging from fieldworkers to philologists, and from

typologists to historical linguists.

For the most part, suggestions to the effect that Riau Indonesian is not a proper language generally involve claims that it is lacking some crucial feature perceived to be a *sine qua non* for true languagehood; typically, such claims invoke some kind of label with intended negative connotations. Following are 12 claims, each of which has been made at some point or another with regard to Riau Indonesian in order to argue that it is not a real language, or at least not an ordinary, run-of-the-mill, or prototypical one:

- (1) Riau Indonesian is just ...
 - Claim 1:* a hoax
 - Claim 2:* a corrupt, broken, imperfect language variety
 - Claim 3:* a language variety spoken only by uneducated people
 - Claim 4:* a language variety without native speakers
 - Claim 5:* a language variety without first-language speakers
 - Claim 6:* a local accent of Indonesian
 - Claim 7:* an artefact of code-switching
 - Claim 8:* a mixed language
 - Claim 9:* foreigner talk
 - Claim 10:* a trade jargon
 - Claim 11:* a pidgin
 - Claim 12:* a creole

While largely independent of each other, the above 12 claims are nevertheless interrelated in various ways, as a result of which some of them sometimes end up being bundled into a single critique. This paper examines the above 12 claims one after another, and argues that each and every one of them is factually wrong.

Claims 1-12 are ordered in very roughly increasing order of seriousness. Claims 1-3 are irrelevant and prejudiced, and it is unfortunate and even a little embarrassing that time and space still need to be wasted in order to dismiss them. Claims 4-12 are at least substantive, however they are all factually wrong — the first eight manifestly so, the last one, being of historical nature, in the absence of any positive corroborating evidence.

It should be noted, though, that with the exception of the first, the above claims should not, even if true, have any bearing on whether Riau Indonesian is a "real language". (Try telling a creolist — cf. Claim 12 — that their object of investigation is not real!) Most commonly, notions of what constitutes a "real language" presuppose a host of normative assumptions about a language's past and present: for example that it should be the result of "normal" transmission down the generations, that it should be the native language of monolinguals — pick your favourite prejudice. But if all of these assumptions were put together, there would be very few proper languages left. Even if one or more of Claims 2 - 12 turned out to be true, Riau Indonesian would still be worthy of serious linguistic investigation. However, as is argued here, there is good reason to reject each and every one of the 12 claims in (1) above.

2.1. "A HOAX"

In response to the publication of the *Economist* article, a linguist writing on the sci.lang blog (<http://groups.google.com/group/sci.lang/msg/f127f1ecf3a5e1d4>) grumbled that "the only way you can experience the real [Riau Indonesian] seems to be to hang around down at the dock with David Gil. Which makes it hard to evaluate his hypothesis". Such

doubts have been expressed in more outspoken terms. At a recent conference, the 12th International Symposium on Malay/Indonesian Linguistics (26-27 June 2008, Leiden, The Netherlands), during a panel discussion, one local linguist said that he could not believe in the existence of Riau Indonesian in the absence of any available published data, while a compatriot of his simply asserted that there is no such thing as Riau Indonesian. An Indonesian linguist in the audience went even further, pronouncing Riau Indonesian to be a "hoax", comparing it to the "Great Eskimo Vocabulary Hoax", which held that the Eskimos had a huge number of different words for 'snow' (Pullum 1989).

The litmus test for any scientific claim is that it be replicable. Well, for anybody wishing to experience Riau Indonesian first hand, directions and travel tips for Riau are readily available in any number of guidebooks. It's easy to get there: that infamous dock, where I collected some of my data, is just half an hour by luxury ferry from the even more luxurious Changi Airport in Singapore, and you're less likely to be mugged there than on most linguists' own university campuses. But the less mobile linguist also has ways in which to independently assess the validity of the data — and there's a lot of it — from the reassuring comfort of his or her desk. When I first started working on Riau Indonesian, I had no research budget or technical assistance; accordingly, the method I developed for collecting naturalistic speech data involved jotting down individual utterances that I heard into a notebook and then entering them into a database. It is such data that is cited in my earlier works on Riau Indonesian. In the meantime, however, my circumstances have greatly improved, and for the last few years the Max Planck Institute for Evolutionary Anthropology has made it possible to employ a team of assistants in Universitas Bung Hatta in Padang, whose task is to transcribe and annotate naturalistic speech in Riau Indonesian, alongside various other dialects and languages. The data, together with data collected from other sites across Indonesia, is housed in a database containing 5 basic fields, providing transliteration, phonetic transcription, interlinear gloss, free translation (into either English or Standard Indonesian) and other comments; the database also provides rich metadata concerning the speakers, the situation in which the recording was made, and so forth, plus links to the original sound files. A subset of the data, including that from Riau Indonesian and closely related languages, is posted on the web, where it is publicly accessible: just go to <http://lingweb.eva.mpg.de/jakarta/data.php>. Linguists are invited to make use of this resource, to reassure themselves that Riau Indonesian is not a hoax, and then hopefully join in the further study of Riau Indonesian and other language varieties.

2.2. "A CORRUPT, BROKEN , IMPERFECT LANGUAGE VARIETY"

Some linguists, accepting that it is not a hoax, are still reluctant to take data from Riau Indonesian seriously, on the grounds that it represents a "corrupt", "broken", or "imperfect" variety of Indonesian. Such attitudes stem from an entrenched tradition of prescriptivism and language engineering which views the standard language as the only legitimate variety and any deviation from it as undesirable. Whatever the possible merits of prescriptivism vis à vis society at large, from the point of view of descriptive linguistics, such attitudes are simply irrelevant. Modern descriptive linguistics takes for granted that, just as all of the world's languages are equally worthy of study, so all varieties of a single language, be they acrolectal or basilectal, are equally deserving of serious scientific investigation. Unfortunately, however, the descriptive approach to linguistics still meets with substantial resistance in many parts of the world, including Indonesia.

Indeed, one might arguably go one step further, turning the tables on the old prejudices, to suggest that, if anything, basilectal varieties of a language are more real than their acrolectal standardized counterparts. Here, briefly, are some reasons why Riau Indonesian might actually claim to be more real than Standard Indonesian. First, Riau Indonesian has native speakers (see Sections 2.4 and 2.5 below), whereas Standard Indonesian, by anybody's account, has none. Secondly, in Riau there are many speakers of Riau Indonesian who have little or no competence in Standard Indonesian, but few or no speakers of Standard Indonesian without complete or near-complete competence in Riau Indonesian. Thirdly, even amongst diglossic individuals with complete mastery of both basilect and acrolect, Riau Indonesian is used with massively greater frequency, and in an overwhelmingly greater variety of situations, than Standard Indonesian. Fourthly, Riau Indonesian shares more typological features with neighboring Southeast Asian languages, while Standard Indonesian exhibits more areally atypical properties resulting from language engineering, much of which involves attempts to mimic perceived characteristics of prestigious western languages. Some examples: (a) Riau Indonesian, like most languages of Southeast Asia, uses the same word for 'and' and 'with', while Standard Indonesian, like most languages of Europe, has distinct words for both functions (Gil 2004a,b); (b) Riau Indonesian, like most languages of Indonesia, has adnominal distributive numerals, formed by reduplication, whereas Standard Indonesian, like English, Dutch and Arabic, has no distributive numerals (Gil 2005b); and (c) Riau Indonesian, like many languages of Southeast Asia, has neither rigid clausal word order nor case marking, whereas Standard Indonesian follows the general European pattern of having at least one of the two — in the case of Standard Indonesian this being rigid clausal word order (Gil 1994, 2005e, 2008a). For these and other reasons, it would seem justified to consider Riau Indonesian (as well as other, similar basilectal varieties of Indonesian) as ontologically prior, and Standard Indonesian as derivative. Indeed, in view of the often inept and linguistically uninformed machinations of the language engineers, it would seem more appropriate to characterize Standard Indonesian as being the "corrupt", "broken" and "imperfect" version of the real language.

2.3. "A LANGUAGE VARIETY SPOKEN ONLY BY UNEDUCATED PEOPLE"

In a related but still distinct complaint to the preceding one, data from Riau Indonesian is sometimes criticized by linguists as being obtained from uneducated, low-class speakers. True, most of the data from Riau Indonesian happens to come from uneducated speakers, but so what? Notwithstanding the provenance of perhaps most of its practitioners, linguistics has never defined itself as the study of middle- or upper-class language.

In fact, in the case of basilectal language varieties such as Riau Indonesian, there are distinct methodological advantages to working with uneducated speakers. Of course, educated persons also speak Riau Indonesian; however, their better mastery of Standard Indonesian means that, when asked to provide judgments in Riau Indonesian, their responses are more likely to be contaminated by their knowledge of the standard language, and their inability — even when willing — to differentiate between the two. For example, in a cross-linguistic experiment currently in progress (see Gil 2007, 2008a for some preliminary results), speakers of Riau Indonesian and other colloquial varieties of Indonesian are asked to judge the availability of various interpretations of sentences involving constructions such as bare, unmarked peripheral non-arguments (eg. *Badut minum buku*, literally 'clown drink book', with the interpretation 'The clown is drinking while reading a book') and non-canonical word orders (eg. *Ayam makan*, literally 'chicken

eat', with the interpretation 'Something is eating the chicken'). Consistently, in Riau Indonesian and other basilectal varieties, uneducated speakers are more willing to accept such interpretations than their highly educated counterparts. As argued in Gil (2005a), this is due, at least in part, to the fact that more highly educated speakers are more likely to be influenced in their judgments by Standard Indonesian, in which such interpretations are not available. What this suggests, then, is that in situations involving diglossia, such as that of Riau Indonesian, uneducated speakers may be better sources of information concerning the language in question, and the data that they provide may in fact be more "real".

2.4. "A LANGUAGE VARIETY WITHOUT NATIVE SPEAKERS"

Adopting a somewhat different tack, it is sometimes claimed that Riau Indonesian is not a real language in the sense that it does not have any native speakers of its own. But this is patently false: most children growing up in Riau speak Riau Indonesian before they reach school age. Moreover, there is no reason to suppose that the facts were significantly different within the lifetime of all current speakers of Riau Indonesian. Of course, many children growing up in Riau may also speak other dialects of Malay/Indonesian and/or other languages before school age. However, multiglossia and multilingualism are the norm in Indonesia as in most other parts of the world, and this is not taken to suggest that any of the dialects or languages involved are any less real for this reason.

2.5. "A LANGUAGE VARIETY WITHOUT FIRST-LANGUAGE SPEAKERS"

In a variant on the preceding claim, it is occasionally suggested that, while young children may indeed be acquiring Riau Indonesian, it is never actually their first language variety: they are invariably acquiring some other dialect or language as their first and hence dominant language variety, and Riau Indonesian only at some subsequent stage of development. Again, this is not the case. To begin with, it is clear that many young children are only acquiring Riau Indonesian; for such children this claim is trivially false. However, as mentioned in the preceding paragraph, there are also many children growing up with multiglossic and/or multilingual competence, and for such children, one or another of their language varieties may indeed be first, in the sense that it is acquired more rapidly and therefore dominates the young child's usage. So which language variety is this? While no careful sociolinguistic studies have been made, impressionistically it would seem that there are a range of situations: while in some cases Riau Indonesian might be the first language, in other cases some other dialect or language would appear to have priority, and in yet other cases it is hard to pick out one of the two or more language varieties as being the dominant one. To summarize, then, it is clear that in very many cases, Riau Indonesian is a first or dominant language, and that in many other cases, it shares this designation alongside one or more other language varieties. Moreover, even in those cases where it comes in second to some other language variety, it still retains its status as a native language acquired in early childhood, and is therefore no less real than the more dominant language variety that the child has acquired.

2.6. "A LOCAL ACCENT OF INDONESIAN"

Riau Indonesian is sometimes characterized as nothing more than Indonesian with a regional accent, reflecting the influence of the indigenous language, Riau Malay, and/or the languages of migrant communities, in particular Minangkabau. It is indeed the case that the pronunciation of Riau Indonesian differs from that of other regional varieties of

Indonesian in ways that reflect the influence of Riau Malay and Minangkabau; some examples of this are presented in Section 3, Table 2 below. However, as commonly understood, the term "accent" refers *only* to pronunciation, whereas the differences between Riau Indonesian and other regional varieties span the entire range of linguistic domains: lexicon, morphosyntax, discourse, and so forth. Thus, Riau Indonesian is not just Indonesian with a Riau accent.

2.7. "AN ARTEFACT OF CODE SWITCHING"

It is occasionally claimed that Riau Indonesian is not a real language but rather a mere artefact of code switching: the speech that results when people mix the languages that they are familiar with, in case at hand Minangkabau, Riau Malay and Indonesian. The claim is a *prima facie* plausible one, given that, throughout Indonesia, as in other countries endowed with comparable linguistic diversity, code switching is so widespread that it is often hard to find a stretch of speech that is completely devoid of code switching, and which may therefore be characterized as being a "pure" instance of some particular language or dialect.

A characteristic feature of code-switching is that it occurs on the fly, reflecting the real-time performance of individual speakers. Accordingly, the product of code switching lacks stability, instead being characterized by a substantial amount of variation. For example, in texts exhibiting code switching, one may encounter the same word once in one language, once in another, without any clear systematic reason. However, Riau Indonesian exhibits little of the instability characteristic of the speech resulting from code switching. For example, if a particular Riau Indonesian word happens to be shared with Minangkabau but not Riau Malay, the corresponding Riau Malay word will never be used; conversely, if another Riau Indonesian word happens to be shared with Riau Malay but not Minangkabau, the corresponding Minangkabau word will never be used — see Section 3 Table 1 for examples. Analogous observations hold also with respect to other linguistic domains such as phonology and morphosyntax. Thus, the stable nature of Riau Indonesian belies the claim that it is a mere artefact of code switching. Moreover, the arbitrary nature of the mixture — as, for example, where one word is shared with Malay but another with Minangkabau — resists any analysis in terms of general principles of code switching, and can only be attributed to a fixed set of conventions part of the competence of speakers of Riau Indonesian.

But there is another, knock-down argument against the characterization of Riau Indonesian as the product of code-switching. In addition to the features that Riau Indonesian shares, in various permutations, with Minangkabau, Riau Malay and/or Standard Indonesian, there are also features that are distinctively Riau Indonesian, that is to say, not present in any of those languages. Examples of such features are presented in Section 3 below. Quite obviously, such features could never appear in speech that is mere code-switching between Minangkabau, Riau Malay and Standard Indonesian.

Although Riau Indonesian is not a product of code switching itself, it must be acknowledged that speakers do frequently code-switch between Riau Indonesian and the languages with which it is in closest contact, Minangkabau, Riau Malay and Standard Indonesian; for examples and discussion of such code switching, see Gil (2004a). However, such code switching is clearly distinguishable from "straight" Riau Indonesian with its conventionalized mixture of features from those three languages. Code switching occurs in well-defined contexts, and necessarily involves speakers who are fluent in all of the respective languages. In contrast, monolingual and monoglossic speakers of Riau

Indonesian will still use whatever features Riau Indonesian shares with Standard Indonesian, Riau Malay and Minangkabau without being able to speak any of those other three languages, thereby showing that Riau Indonesian is not a mere artefact of code switching.

2.8. "A MIXED LANGUAGE"

If the mixture of languages manifest in Riau Indonesian is a stable one, then perhaps — so it has been suggested — Riau Indonesian is a mixed language: the kind of language that results from the conventionalization over time of code switching. Other cases of mixed languages involving Malay/Indonesian have been proposed, including Sri Lankan Malay, a mixture of Malay and Tamil (Slomanson to appear), Semarangan, a mixture of Malay and Javanese (Tadmor 2005), and Steurtjestaal (van Rheedeen 1999), a mixture of Malay, various regional languages and Dutch. In fact, in the case of Riau Indonesian, at least two different suggestions have been made with regard to the identity of the source languages: one that it is a mixture of Indonesian and Minangkabau, the other that it is a mixture of Indonesian and Malay, where the terms "Indonesian and "Malay" apparently refer to prestige versions of the two languages, either the respective standard varieties, or alternatively the colloquial varieties spoken in the respective capital cities, Jakarta and Kuala Lumpur.

Obviously, since Riau Indonesian is in close contact with Minangkabau, it has taken on a number of linguistic features from Minangkabau; several examples of such features are provided in the next section. Similarly, given that Sumatra, on which Riau Indonesian is spoken is situated in-between Java, where Jakarta is located, and the Malay peninsula, where Kuala Lumpur is located, it is hardly surprising that, with respect to a wide range of linguistic features, Riau Indonesian occupies an intermediate position between Jakarta Indonesian and Kuala Lumpur Malay, or between Standard Indonesian and Standard Malay; see Section 4 for further discussion.

Nevertheless, Riau Indonesian clearly lacks certain characteristics that are generally associated with mixed languages. First, in mixed languages, the two source languages are generally manifest in different linguistic domains; most commonly, one is dominant in the lexicon while the other constitutes the lion's share of the morphosyntax. Thus, for example, in Sri Lankan Malay, the lexicon is predominantly Malay whereas the morphosyntax is largely Tamil. However, in Riau Indonesian, features from each of the two putative source languages are spread across all of the linguistic domains, including the lexicon and the morphosyntax. This is shown clearly in Section 3 below, in Tables 1-4 and subsequent discussion. Secondly, in mixed languages, the two source languages are typically very different from each other, lexically and grammatically; for the most part it is easy to tell which bits of the mixed language come from one source language and which bits from the other one. However in Riau Indonesian, the would-be source languages are so similar to each other that it is often very hard to tell whether a particular word or construction comes from Indonesian or from the other language, be it Riau Malay or Minangkabau. Thus, there would seem to be little reason to characterize Riau Indonesian as a mixed language.

2.9. "FOREIGNER TALK"

Occasionally, it is suggested that my picture of Riau Indonesian is coloured by my obvious outsider status, and that what I am describing is some kind of foreigner talk. This is not nearly as silly a suggestion as it might sound. Indonesians modulate their

speech in many ways depending on the ethnicity of their interlocutor. To begin, even within a well-defined language variety, such as Riau Indonesian, there are different terms of address for Westerners, Chinese, Malays from Malaysia, Javanese, and indigenous people. For example, when addressing a male speaker of roughly similar age, a speaker might use *ster*, *pek*, *cik*, *mas* and *bang* respectively. It is also customary to adopt the interlocutor's real or perceived accent; for example, whereas Riau Indonesian has no lexical stress (Gil 2003b, 2006a), people sometimes adopt penultimate stress when speaking to westerners, or final stress when talking to Chinese: this is not considered to be condescending or insulting as it might be in many other parts of the world. Indeed, since the most typical venue for communicating with foreigners is in the marketplace, Bazaar Malay, mentioned in Section 2.10 below, has come to assume the role of a conventionalized foreigner talk. Thus, Riau, like much of the region, is rife with foreigner talk, and it is therefore incumbent on researchers to make sure that their data is not contaminated by this phenomenon.

Still, anybody with a modicum of familiarity with the linguistic landscape of a particular region of Indonesia can easily distinguish foreigner talk from the "real thing". In particular, in my own sojourns in Riau, it is easy to tell when the occasional stranger is addressing me in some kind of foreigner talk. In general, people who know me would never use foreigner talk with me, since to do so might be construed as an impolite expression of social exclusion. In any case, a large proportion of my data is based on conversations amongst locals in which I was not involved. Thus, Riau Indonesian is, quite emphatically, not foreigner talk.

2.10. "A TRADE JARGON"

Alternatively, Riau Indonesian is sometimes taken to be nothing more than just a trade jargon, the language of marketplaces and merchant vessels. Sociolinguistically, however, Riau Indonesian is nothing of the sort. Unlike trade jargons, it has native speakers; unlike trade jargons, it is used in a wide range of everyday situations, in home, school, mosque, football field — just about everywhere. In fact, throughout large parts of the archipelago, including Riau, there is a trade jargon with a traditional name: *Bazaar Malay*. But Riau Indonesian is clearly structurally distinct from Bazaar Malay. To cite just two differences: unlike in Bazaar Malay, genitives are usually postnominal; again unlike in Bazaar Malay, forms such as *di-* and *N-* are used to mark generalized passive and active voice (Gil 2002b). Thus, Riau Indonesian is not Bazaar Malay, nor is it, in any sense of the term, some kind of trade jargon.

2.11. "A PIDGIN"

Related to the two previous claims, it is occasionally suggested that Riau Indonesian might be a pidgin. Since one of the hallmarks of a pidgin is a radically reduced grammar, such suggestions are generally motivated by my characterization of Riau Indonesian as having a very simple grammar. Be that as it may, Riau Indonesian fails to meet the profile of a typical pidgin on multiple other grounds. Structurally, pidgins generally also have a small lexicon, but that of Riau Indonesian is in the same ballpark as those of most other languages. Moreover, pidgins are typically unstable, exhibiting lots of variation, whereas Riau Indonesian, as argued in Section 2.7 above, is as stable as any other language. Sociolinguistically, pidgins lack native speakers, whereas Riau Indonesian, as observed in Section 2.4 above, has lots. In addition, pidgins are typically used in a limited set of communicative contexts, whereas Riau Indonesian, as noted above, is used

in a wide range of contexts. So for these reasons, at least, Riau Indonesian is clearly not a pidgin.

2.12. "A CREOLE"

If not a pidgin, then perhaps Riau Indonesian is a creole. Other versions of Malay/Indonesian have been argued to be creoles, or the descendants thereof, such as Peranakan Malay (Pakir 1986), Jakarta Indonesian (Tadmor 2007), and several eastern Indonesian varieties (Pauw 2008). Moreover, as noted in Gil (2001a), the typological resemblance between Riau Indonesian and other "classical" creole languages is often cited as evidence that Riau Indonesian is also a creole language. This argument is made most forcefully by McWhorter (2001, 2005), who maintains that the apparent absence of grammatical complexity in Riau Indonesian could only have resulted from the kind of abrupt break in transmission that is criterial of creole languages. Moreover, in response to the observation that other colloquial varieties of Malay/Indonesian share the typological profile of Riau Indonesian, McWhorter suggests that the creole label may in fact apply to most or all spoken varieties of Malay/Indonesian.

However, even if some varieties of Malay and Indonesian are indeed creoles, there is no independent evidence to the effect that, at some stage in the past, there was an abrupt break in transmission resulting in the radical restructuring of a language ancestral to Riau Indonesian. For McWhorter (2008), the absence of such independent evidence does not matter: on the basis of other cases where such evidence is available, he argues that the absence of grammatical complexity characteristic of Riau Indonesian could *only* have arisen as the result of an abrupt break in transmission, and therefore we are entitled to assume that this happened even in cases where the historical evidence is not available. However, while it may be true that radical restructuring results in a typological profile similar to Riau Indonesian, it is less obvious that the absence of grammatical complexity can *only* result from such an abrupt break in language transmission. Under an alternative scenario put forward in Gil (2001a), the overall absence of grammatical complexity in Riau Indonesian might more appropriately be construed as an accidental confluence of a number of independent areally-motivated diachronic processes of simplification in independent components of the grammar, each of which is attested elsewhere in the world, in clearly non-creole languages. Thus, the typological profile of Riau Indonesian cannot be invoked in support of its characterization as a putative creole language.

Moreover, Riau Indonesian lacks one important sociolinguistic property typical of creole languages. Speakers of creole languages are invariably of an ethnicity distinct from that of the speakers of the lexifier languages. This is clearly the case with respect to the classical plantation creoles that developed far away from their European lexifier languages: one would not expect to find an English-based creole cropping up amongst an ethnically English community in England. But this is also equally the case with regard to the most reasonable candidates for Malay/Indonesian-based creoles, none of which are spoken by ethnic Malays. Thus, for example, Peranakan Malay is spoken by the Peranakan Chinese, Jakarta Indonesian is spoken by an urban population consisting of Betawi, Sundanese and Javanese and a medley of other ethnicities, while eastern Indonesian varieties such as Ambon, Kupang and Papuan Malay are each spoken by people of many diverse ethnicities. In contrast, however, Riau Indonesian is spoken in a region where the indigenous population shares the ethnicity of the would-be lexifier language, namely Malay. Indeed, the region where Riau Indonesian is spoken has been Malay-populated and Malay-speaking for the last 1500 or 2000 years at least.

Admittedly, immigrants have been flocking into Riau for many hundreds of years, but they have always adopted the local language the same way as immigrants to England have always adopted English, without, usually, creating a creole. Thus, it is very implausible to assume that Malay speakers in an indigenously Malay region such as Riau would, suddenly and for no apparent reason, restructure their language in order to create a new Malay/Indonesian-lexifier creole language.

This is not to deny that much of the structural simplicity of Riau Indonesian may result from processes pertaining to language contact; indeed, in more recent work, McWhorter (2006) offers a plausible characterization of Malay/Indonesian as a Non-Hybridized Conventionalized Second Language (NCSL) — a language of a type, including other major world languages such as English, Persian, Arabic and Mandarin, where rampant second-language acquisition results in a structure that is simpler than that of the language's close non-NCSL relatives. This is also not to dismiss the possibility that in the more distant part, some ancestor of Malay/Indonesian or even proto-Malayic may have been a creole language; thus, Donohue and Denham (to appear), citing a range of grammatical, archaeological and genetic evidence, suggest that the original expansion of Malayo-Polynesian languages south into the Indonesian archipelago may have been in the form of a handful of original creole languages from which most or all of the extant Malayo-Polynesian languages are descended. However, whatever the merits of their proposal, Donohue and Denham's Malayo-Polynesian creoles pertain to time depths that are just too far in the past to be of relevance to contemporary Riau Indonesian and whatever features may distinguish it from other varieties of Malay/Indonesian or from other languages elsewhere.

2.13. INTERIM SUMMARY

We have now examined and hopefully put to rest, in turn, each of the 12 claims in (1) concerning the nature of Riau Indonesian. The conclusion, then, is there is nothing "wrong" with Riau Indonesian and that it is a real language, or at least as real as any other of the world's thousands of languages. Of course, there would have been nothing "wrong" with Riau Indonesian even if it had turned out to be, say, an artefact of code-switching, or a creole. More generally, the preceding discussion suggests that the historical and sociological labels that we use to describe languages and linguistic behaviour are not only too laden with undesirable connotations but also insufficient to capture the diversity of situations in which language occurs — the reader may well be left at this point wondering "what kind of language" Riau Indonesian actually is. Arguing what something *is not*, as we have done so far, is a rather unsatisfying chore; it is far more interesting to assert what something *is* — what it is like, and how it resembles and differs from other related entities. So in the next two sections we turn to consider the nature of Riau Indonesian, focusing on its relationships with the languages with which it is in closest contact, in sociological space in Section 3, and in geographical space in Section 4.

3. RIAU INDONESIAN AND ITS SUBSTRATE AND SUPERSTRATE LANGUAGES

Although, as argued in Section 2.12 above, Riau Indonesian is not a creole language, it makes sense to borrow from creolistic terminology in order to talk about its substrate and superstrate languages. The two main substrate languages are Riau Malay, the indigenous language of Riau, and Minangkabau, the language spoken by the largest number of

immigrants to Riau. The superstrate language of Riau Indonesian is of course Standard Indonesian, with respect to which it stands in a basilect-to-acrolect relationship.

One obvious difference between Riau Indonesian and most creole languages is that for creoles, the substrate and superstrate languages are generally very different from each other; moreover, they are manifest in different linguistic domains — the substrate largely in the phonology and morphosyntax, the superstrate mostly in the lexicon. In contrast, for Riau Indonesian, the substrate and superstrate languages are closely related to each other, and, ipso facto, also to Riau Indonesian; as a result, with respect to numerous features, Riau Indonesian is indistinguishable from its substrates and its superstrate. Thus, whereas in creole studies, the creolist's challenge is to find features of the creole language that originate in its substrate, Riau Indonesian turns the tables: here the challenge is to identify features that are *not* attributable to its substrate, or, for that matter, its very similar superstrate. In fact, it is precisely these features that attest to the reality of Riau Indonesian as distinct from its substrate and superstrate contact languages.

3.1. SHARED AND DISTINCTIVE FEATURES

Some data reflecting the relationship between Riau Indonesian and its substrate and superstrate languages is presented in Tables 1-4. Tables 1-4 represent, respectively, four domains of linguistic structure: lexicon, phonology, morphosyntactic matter (involving the specific phonological forms of grammatical items), and morphosyntactic pattern (involving the abstract structures into which various items may enter). The tables underscore the extent of the similarity between Riau Indonesian and its two substrate languages Minangkabau and Riau Malay, as well as its superstrate Standard Indonesian; since Riau Malay exhibits a substantial degree of geographical variation, one particular subdialect is chosen, namely Siak Malay, spoken in the lower Siak river basin, some aspects of which are described in Gil (2001a, 2002a, 2004a). Each table contains eight rows corresponding to the eight logically possible patterns of identity and non-identity between Riau Indonesian and the three languages in question, abbreviated *Min* (Minangkabau), *SkM* (Siak Malay) and *StI* (Standard Indonesian); each row presents an item in Riau Indonesian exemplifying that particular pattern of identity. Thus, row 1 presents a Riau Indonesian item shared by all three languages, row 2 a Riau Indonesian item shared by Minangkabau and Siak Malay but not Standard Indonesian, and so on all the way down to row 8 presenting a Riau Indonesian item present in none of the three languages.

1	Min	SkM	StI	<i>makan</i> 'eat'
2	Min	SkM	-	<i>tengok</i> 'look'
3	Min	-	StI	<i>kuap</i> 'yawn'
4	Min	-	-	<i>hijau</i> 'grue'
5	-	SkM	StI	<i>ikan</i> 'fish'
6	-	SkM	-	<i>pompong</i> 'wooden boat with engine'
7	-	-	StI	<i>celana</i> 'trousers'
8	-	-	-	<i>kasi</i> 'give'

Table 1: *Shared and Distinctive Lexical Features*

1	Min	SkM	StI	no word-final palatals (in ordinary words)
2	Min	SkM	-	no lexical stress
3	Min	-	StI	no fronting of /a/ before coronal consonants word-finally
4	Min	-	-	no phonemic /ə/
5	-	SkM	StI	word-final /t/ (in ordinary words)
6	-	SkM	-	no high vowel in word-final closed syllable (in ordinary words)
7	-	-	StI	word-final /r/ (in ordinary words)
8	-	-	-	absence of high vowel in word-final closed syllable overridden by identical vowel harmony (in ordinary words)

Table 2: *Shared and Distinctive Phonological Features*

1	Min	SkM	StI	patient-orientation <i>di-</i>
2	Min	SkM	-	agent-orientation <i>N-</i>
3	Min	-	StI	2PL pronoun <i>kalian</i>
4	Min	-	-	negative <i>ndak</i>
5	-	SkM	StI	proximal future <i>nanti</i>
6	-	SkM	-	intensifier <i>betul</i>
7	-	-	StI	temporal interrogative <i>kapan</i>
8	-	-	-	applicative <i>-in</i>

Table 3: *Shared and Distinctive Morphosyntactic Matter Features*

1	Min	SkM	StI	optional numeral classifiers
2	Min	SkM	-	negative polarity particle
3	Min	-	StI	regular negation for 'want'
4	Min	-	-	sentence final particle 'one'
5	-	SkM	StI	single agent-orientation prefix
6	-	SkM	-	inclusory plural construction
7	-	-	StI	non-alternating associative enclitic
8	-	-	-	associative disjunction construction

Table 4: *Shared and Distinctive Morphosyntactic Pattern Features*

Additional clarificatory information regarding the contents of Tables 1-4 and the specific items therein is presented in the Appendix.

3.2. WHAT THESE FEATURES MEAN

What do tables 1-4 tell us? The most obvious fact is that, in each table, each of the eight rows is exemplified with some linguistic item from Riau Indonesian; that is to say, within each linguistic domain, all eight logically possible patterns of shared features with respect to the three contact languages are actually attested. It should be acknowledged, however, that Tables 1-4 say nothing about the relative frequencies of these different patterns, which in fact are far from equal. Specifically, within each table, the pattern in row 1 is

overwhelmingly more common than any of the others. In other words, the most common state of affairs is one in which linguistic features are shared by Riau Indonesian, Minangkabau, Siak Malay and Standard Indonesian, reflecting the fact that these four language varieties are very closely related to each other.

In each table, rows 1,2,3,4 represent cases of identity between Riau Indonesian and Minangkabau, rows 1,2,5,6 identity between Riau Indonesian and Siak Malay, and rows 1,3,5,7 identity between Riau Indonesian and Standard Indonesian. Thus, in four out of the eight rows, the Riau Indonesian item is identical to more than one of the three contact languages, in which case it is difficult to relate the Riau Indonesian item unequivocally to one specific contact language to the exclusion of one or more of the others. However, in three other rows, the Riau Indonesian item is identical to exactly one of the three contact languages, and in these cases, a connection to the respective language is indicated. Thus, row 4 presents items which Riau Indonesian shares with Minangkabau to the exclusion of Siak Malay and Standard Indonesian; these items thus constitute plausible candidates for a Minangkabau substrate influence on Riau Indonesian, though other potential scenarios are also conceivable, such as, for example, borrowing from a colloquial variety of Indonesian similar to Riau Indonesian into Minangkabau. Similarly, row 6 presents items which Riau Indonesian shares with Siak Malay to the exclusion of Minangkabau and Standard Indonesian; these items constitute likely candidates for a Malay substrate influence on Riau Indonesian, though, once again, alternative scenarios need also to be considered. Conversely, row 7 presents items which Riau Indonesian shares with Standard Indonesian to the exclusion of Minangkabau and Siak Malay; these items constitute possible candidates for a Standard Indonesian superstrate influence on Riau Indonesian, though, given the derivative nature of the standard language argued for in Section 2.2 above, it is perhaps more likely that such items entered into Standard Indonesian from some colloquial variety of Malay or Indonesian resembling Riau Indonesian. Finally, row 8 presents items with respect to which Riau Indonesian differs from all three contact languages, underscoring the distinctive nature of Riau Indonesian with respect to those three language varieties.

Tables 1-4 provide further evidence against some of the claims about Riau Indonesian listed in (1) and dealt with in Section 2 above, specifically, those pertaining to the relationship between Riau Indonesian and its contact languages. To begin with, the fact that in each of the four tables all of the rows are instantiated with items from Riau Indonesian shows that Riau Indonesian is not a mixed language as per Claim 8, or a creole as per Claim 12, since, if it were, then one would expect the lexicon to come mostly from one language and the grammar mostly from another.

The same fact also shows that Riau Indonesian is not a mere artefact of code switching as per Claim 7. Consider, for example, Table 1, rows 2 and 7. Why does Riau Indonesian share the word for 'look' with Minangkabau and Siak Malay but the word for 'trousers' with Standard Indonesian? If Riau Indonesian were the result of code-switching, then one would also expect the Standard Indonesian word for 'look', and the Minangkabau and Siak Malay words for 'trousers' to make occasional appearances — but they do not. In view of this, the only way to uphold a code-switching analysis is to formulate a set of general principles that would explain, among other things, why 'look' always comes from Minangkabau and Siak Malay but 'trousers' always from Standard Indonesian. More generally, such principles would have to account for all of the data in Tables 1-4, and indeed for the entirety of the lexicon, phonology and morphosyntax of

Riau Indonesian, explaining why speakers consistently chose the items that they do while rejecting other available items from other languages. However, the diversity of the data in Tables 1-4 suggests that no such set of principles is ever likely to be formulated. In other words, whether a given Riau Indonesian item is the same as its counterparts in Minangkabau, Siak Malay and/or Standard Indonesian is an arbitrary and unpredictable property of Riau Indonesian, one that must be attributed to the native speakers' knowledge of their language, and therefore explicitly represented in the lexicon or grammar of Riau Indonesian. In other words, Riau Indonesian is not a mere product of code switching but rather a proper language with its own unique character.

Finally, and most strikingly, within all four tables, the items in row 8 represent a small residue of features of Riau Indonesian that differ from Minangkabau, Siak Malay and Standard Indonesian. These features prove that Riau Indonesian is not just the result of these three languages coming together, as would be the case if it were the product of code switching as per Claim 7, or a mixed language as per Claim 8. The question arises where these features come from; and there is apparently no single answer. Some most likely originate in other adstrate dialects with which Riau Indonesian is in contact, such as *kasi* 'give' from Bazaar Malay in Table 1, or applicative *-in* from Jakarta Indonesian in Table 3. Others, however, are probably most appropriately analyzed as the results of internal developments within Riau Indonesian, such as the specific vowel harmony constraint in Table 2. Whatever their origins, though, the presence of these items provide the clinching argument for the distinctiveness of Riau Indonesian as a language in its own right, with its unique stable and well-defined lexicon and grammar.

4. RIAU INDONESIAN AND ITS NEIGHBORING LANGUAGES

Having examined the sociolinguistic circumstances of Riau Indonesian, the question arises whether there is anything special about the sociolinguistic landscape of Riau, or whether, alternatively, Riau Indonesian has its counterparts in other colloquial varieties of Indonesian spoken in other parts of the vast archipelago. The answer to this question is, most emphatically, the latter.

4.1. OTHER REGIONAL VARIETIES OF COLLOQUIAL INDONESIAN

Although every place in Indonesia has its own unique circumstances, there are sufficiently many commonalities across the diverse regions of Indonesia for colloquial varieties of Indonesian, corresponding to Riau Indonesian, to have developed in most or all of them. Sociolinguistically, Riau Indonesian is anything but exceptional, except in the trivial sense that each and every language and dialect has its own unique properties. Just as there is a Riau Indonesian, so there are varieties of Indonesian associated with most other regions of Indonesia.

What is striking is how little attention such regional varieties of Indonesian have attracted from the linguistic community. For some reason, the eastern parts of the archipelago have fared relatively better, and in recent decades there have been a number of studies of varieties such as Manado Malay (Prentice 1994, Stoel 2005), North Maluku Malay (Voorhoeve 1983, Taylor 1983), Ambon Malay (Collins 1980, van Minde 1997), Kupang Malay (Steinhauer 1983), Papuan Malay (Donohue and Sawaki 2007, Donohue to appear), and one or two others. (Although such eastern varieties are traditionally referred to as dialects of Malay, recent decades have witnessed a terminological shift whereby current speakers of these varieties are more likely to refer to them as Indonesian;

and indeed, with respect to their sociolinguistic functions, they closely resemble varieties of colloquial Indonesian spoken in other parts of the country.) In contrast, with respect to colloquial varieties of Indonesian, the center and west of the country, totaling perhaps close to two hundred million speakers, are largely terra incognita. One obvious exception is the colloquial Indonesian of the capital city Jakarta, which has been the object of a significant number of recent studies (Wouk 1989, 1999, Cole and Hermon 2005, Tjung 2006, Sneddon 2006, Gil 2006b, 2008b and others). However, to the best of my knowledge, other than Jakarta Indonesian, there have been no published studies of any central or western varieties of colloquial Indonesian other than my own on Riau Indonesian. As a result, we know next to nothing about the Indonesian spoken in most of Sumatra, Borneo, Java and Bali, including major metropolitan centers such as Medan, Pontianak, Surabaya and numerous others. Indeed, in those cases where, as in Riau, the local language is a variety of Malay, there are often studies available on the local Malay — but never, as far as I am aware, anything on the local Indonesian. For example, for the province of Jambi, situated just to the south of Riau, a recent NSF-funded project has greatly increased our knowledge of the Malay dialects spoken in the city of Jambi and some surrounding villages (Durvasula 2008, Cole, Hermon and Yanti 2008, Yanti in preparation), while contributing little or no information on the Indonesian spoken in the same places. Of course, part of the reason for this systematic lacuna is that the same prejudices discussed in Section 2 of this paper in the context of Riau Indonesian are equally applicable to other regions. Accordingly, linguists are often not aware of the existence of local varieties of Indonesian, or if they are, they are not interested or otherwise dissuaded from studying them.

On the basis of my own observations from years of traveling around Indonesia, it is my impression that most regions of the country have a distinctive variety of colloquial Indonesian associated with them, regardless of whether the local regional language also happens to be a dialect of Malay. Perhaps not everywhere; there may well be some places in deepest Java or New Guinea where Indonesian has not (yet) made significant inroads, but these would be the exceptions that prove the rule. Of course, each and every regional variety of colloquial Indonesian has its own particular sociolinguistic circumstances. In principle, then, the arguments put forward in Section 2 dismissing the 12 claims in (1) regarding Riau Indonesian need to be applied, again and again, to each and every regional variety of Indonesian; and there is no guarantee that they will always be of equal appropriateness. Just because Riau Indonesian is a stable language variety and not an artefact of code switching, one cannot presuppose that the same is true also for, say, the Indonesian spoken in the province of Kalimantan Barat ('West Kalimantan'); this has to be demonstrated specifically for Kalimantan Barat Indonesian. Similarly, just because there is no evidence that Riau Indonesian is a creole, this does not mean that other varieties of colloquial Indonesian are not creoles; indeed, as pointed out in Section 2.12 above, there is evidence that some probably are. A huge amount of work on the sociolinguistics of colloquial Indonesian dialects is crying out to be done. Still, from my own observations, it would seem that in broad sociolinguistic outlines, Riau Indonesian is anything but exceptional, and that most regions of Indonesia also have their own distinctive varieties of colloquial Indonesian, possessing native speakers, endowed with stable lexicons and grammars, and used in a wide range of situations.

4.2. ISOGLOSSES OF COLLOQUIAL INDONESIAN

The existence of sociolinguistically similar varieties of colloquial Indonesian in other

regions makes it possible to examine Riau Indonesian in geographical perspective, comparing it to neighboring varieties of colloquial Indonesian. For the last several years, I have been engaged in collecting data and constructing dialect maps for colloquial Indonesian (some preliminary results were presented in Gil 2003a). The maps that result from this work look just like ordinary dialect maps, with neighboring dialects tending to exhibiting more similarities than dialects further apart from each other, and isoglosses bundling together to demarcate distinct dialects.

The maps for colloquial Indonesian can be superimposed on the corresponding maps for the regional languages of Indonesia, resulting in a multi-dimensional picture of linguistic reality, where horizontal and vertical axes on the page represent geographical space, and up and out of the page represents sociolinguistic space and, in particular, the distinction between regional languages and the dialects of Indonesian spoken in the same places. For example, a map showing the relative order of nouns and genitives (attributive possessors) for colloquial Indonesian can be superimposed on a corresponding map for the regional languages of Indonesia, such as that derivable from the world map in Dryer (2005). In broad outline, both maps exhibit a similar pattern, with noun-genitive order the rule in the west and genitive-noun order prevalent in the east; this similarity is due in large part to the substrate influence of local languages on the coextensive local varieties of colloquial Indonesian. However, there are also instances of mismatches, such as in northern Sulawesi, where local languages such as Tondano have noun-genitive order whereas the local Manado Malay has genitive-noun order; such mismatches attest to the reality of regional dialects of colloquial Indonesian as autonomous language varieties and not mere relexifications of local languages. (In the case at hand, the genitive-noun order of Manado Malay is one of a large number of features which suggests that the Malay/Indonesian language came to Manado as the result of a back migration from Ternate, located further to the east, where genitive-noun order is widespread in both Malay/Indonesian and the local languages — see Prentice 1994, Paauw 2008.)

Consideration of Riau Indonesian in its geographical context provides yet additional evidence for its reality, and against some of the claims cited in (1) above. Figure 1 presents an idealized map showing Riau in relationship to four other locations strung out on a very rough north-south axis: to the north, Kuala Lumpur, the capital city of Malaysia, located just across the straits of Malacca on the Asian mainland; and to the south, Palembang, the capital city of Sumatra Selatan ('South Sumatra') province, Kalianda, a small town in southern Lampung province, and Jakarta, the Indonesian capital, situated across the Sunda straits on the island of Java. Figure 1 plots the distribution on the idealized map of 8 linguistic features: 4 phonological ones in the middle column, and 4 morphosyntactic ones in the right column. For Kuala Lumpur, these features are plotted with respect to the local variety of colloquial Malay; for Riau with respect to Riau Indonesian, and for the remaining three points with respect to the local varieties of Indonesian corresponding sociolinguistically to Riau Indonesian. Each of the 8 linguistic features distinguishes Kuala Lumpur Malay, at the top of the map, from Jakarta Indonesian at the bottom; however, as evident from their values at the three intermediate locations, Riau, Palembang and Kalianda, they do so in different ways. Specifically, these 8 features define 4 isoglosses, numbered 1-4 in Figure 1, falling in each of the 4 logically possible cut-off points defined by the five geographical locations. The features are all binary, with *yes* and *no* values characterizing the locations above or below the isogloss as indicated. Further explanation of the particular features referred to

in Figure 1 is given in the Appendix.

Kuala Lumpur					
1	-----	word-final /r/	no yes	negative <i>nggak</i>	no yes
Riau					
2	-----	word-final /əC/	no yes	deictic <i>no</i>	no yes
Palembang					
3	-----	word-final /k/ > [ʔ]	no yes	negative <i>tak</i>	yes no
Kalianda					
4	-----	word-final /a/ > [e]	no yes	negative <i>kagak</i>	no yes
Jakarta					

Figure 1: *Some Malay/Indonesian Isoglosses between Kuala Lumpur and Jakarta*

Focusing on Riau Indonesian, Figure 1 suggests that it occupies an intermediate position between Kuala Lumpur Malay and Jakarta Indonesian. Specifically, whereas with respect to the two features constituting isogloss 1, Riau Indonesian resembles Jakarta Indonesian, with respect to the six features constituting isoglosses 2, 3 and 4, Riau Indonesian resembles instead Kuala Lumpur Malay. It is facts such as these that give rise to the impression, cited in Section 2.8 above, that Riau Indonesian might be a mixture of Indonesian (as spoken in Jakarta) and Malay (as spoken in Kuala Lumpur). But this is not really the right conclusion to be drawn from the map. The choice of Kuala Lumpur and Jakarta as beginning and end points is arbitrary; any number of other locations might have been chosen instead, pointing towards different — and equally inappropriate — characterizations of Riau Indonesian in terms of language mixture. In truth, any language or dialect can be characterized as occupying an intermediate position between some language or dialect at an arbitrary distance on one side of it and some other language or dialect at an arbitrary distance on the other side, but such a characterization contributes little towards a better understanding of the nature of the language in question.

A more appropriate conclusion to be drawn from Figure 1 is that dialect maps for colloquial varieties of Indonesian look much like similar maps for other dialects and languages elsewhere. As noted previously, dialects and languages in general tend to bear a closer resemblance to their neighbors than to more distant languages, and Figure 1 suggests that the same is true also for varieties of colloquial Indonesian. As evident from Figure 1, Riau Indonesian shares more features with Palembang Indonesian than it does with the Indonesian spoken in Kalianda, and more with Kalianda Indonesian than with the Indonesian spoken in Jakarta. Moreover, this is true independently of the regional languages which constitute the substrates for the respective varieties of colloquial Indonesian. Admittedly, the substrate for Palembang Indonesian, namely Palembang Malay, is closer to Riau Indonesian than is the substrate for Kalianda Indonesian, a southern dialect of Lampung; this is doubtlessly a large part of the reason why Palembang Indonesian is more similar to Riau Indonesian than Kalianda Indonesian is. However, moving down the map, the Lampung substrate for Kalianda Indonesian is no closer to Riau Indonesian than are the substrate languages for Jakarta Indonesian, primarily Sundanese and Javanese; and yet — as indicated by isogloss 4 — the Indonesian in Kalianda still shares features with Riau Indonesian to the exclusion of Jakarta Indonesian. For such features, then, horizontal contact between a language and its neighbors trumps vertical contact between a language and its substrates. In this particular case, geographical proximity also wins out over the massive onslaught of Jakarta Indonesian across the archipelago via the mass media, which is responsible for the spread of Jakarta Indonesian as a colloquial lingua franca throughout the country. Even in this day and age, colloquial Indonesian language change is carried by slow busses along the narrow, winding, bumpy roads of Sumatra at least as much as by satellite television broadcasts of music, soap operas, and other staples of popular culture.

Thus, what Figure 1 shows is that whether Riau Indonesian possesses a particular feature is determined not just by its position in sociolinguistic space and the influence of substrate and superstrate languages, but also by its location in geographical space and contact with neighboring varieties of colloquial Indonesian. This in turn provides yet further evidence — if such were still needed — that Riau Indonesian and other varieties of colloquial Indonesian are just ordinary dialects, which, like other dialects and languages in other parts of the world, display the kind of geographical patterns that constitute the traditional subject matter of dialect and language atlases.

4.3. BUT IS RIAU INDONESIAN A GEOGRAPHICALLY COHERENT ENTITY?

So geographical considerations join forces with a host of other reasons for viewing Riau Indonesian as a bona fide language or dialect just like so many thousands of others. Yet having dismissed, in the course of this paper, all the proposed reasons for doubting the reality of Riau Indonesian, it must be acknowledged that geography also provides one good reason for questioning the identity of Riau Indonesian defined as "the variety of colloquial Indonesian spoken in Riau". At issue is whether the colloquial Indonesian in one part of Riau is, on the one hand, sufficiently similar to that in all other parts of Riau, and on the other, sufficiently different from that in all other places outside of Riau, to justify talking of a Riau Indonesian, as that dialect associated with all and only the particular geographical region known as Riau.

In one trivial sense, the answer is almost preordained to be negative: this is because the actual location of "Riau" has been moving back and forth at the whims of Indonesian officialdom. In the past, the term Riau referred to a historically and culturally prominent

archipelago, consisting of Bintan, Batam, Karimun, Singkep, and many other smaller islands located to the south of the Malay peninsula and to the east of central Sumatra. When Indonesia gained its independence in 1945, the islands became part of the province of Sumatra Tengah ('Central Sumatra'), but in 1957, this province was divided into three new provinces, one of which was named Riau. However, in addition to the Riau archipelago, the newly created province also contained a largish chunk of the central-eastern Sumatran mainland (including among others, the Siak river basin, where Siak Malay, referred to in Section 3, is spoken). Since this was still the state of affairs when I started working in the region, it seemed reasonable to define Riau Indonesian as the variety of colloquial Indonesian spoken in what was then the province of Riau. However, in 2004, the province of Riau was divided into two, with what was historically the Riau archipelago becoming the province of Riau Kepulauan ('Riau archipelago'), and the remainder, comprising the mainland part plus a few neighboring islands, retaining the simple appellation Riau. So now, what I've been calling Riau Indonesian is actually the Indonesian spoken in two provinces, called Riau and Riau Kepulauan. Of course, this kind of toponymic ping-pong has no effect on the actual language situation; however, it underscores the arbitrariness of political boundaries and serves as a timely reminder that linguists should not expect such artificial lines on the map to provide a reliable indication of linguistic boundaries.

At present, nowhere near enough is known about the local varieties of colloquial Indonesian to be able to determine whether there is any geographical justification to talk of a Riau Indonesian as a well-defined dialect distinct from other neighboring and equally well-defined dialects. What is clear is that the combined linguistic boundaries of the two Riau provinces do correspond, at least roughly, to some linguistic isoglosses. For example, the border between Riau and neighboring Sumatra Barat ('West Sumatra') province also approximates the isogloss separating, among others, the words for 'meet', *jumpa* in Riau and *ketemu* in Sumatra Barat, and for 'put on (clothes)', *pasang* in Riau and *pakai* in Sumatra Barat. (Presumably, in this case, the similarity between administrative and linguistic boundaries is due to the fact that both roughly follow the watershed defined by the mountain range that extends the length of the island.) However, it is also clear that there is some internal variation within the two provinces sharing the name of Riau. For example, the map similar in Figure 1 could be further elaborated in order to distinguish between coastal and interior Riau. Such a map would contain an additional isogloss grouping coastal Riau with Kuala Lumpur to the north, and interior Riau with Palembang, Kalianda and Jakarta to the south; some examples of features instantiating this isogloss and thereby cross-cutting Riau include the words for 'cuttlefish', *sotong* to the north and *cumi* (or *cumi-cumi*) to the south, and for 'anchovy', *bilis* to the north and *teri* to the south. Thus, as presently available, the facts are equally consistent with the existence of a geographically coherent Riau dialect of Indonesian, and alternatively with a state of affairs in which the Indonesian spoken in Riau does not constitute a well-defined dialect, in which latter case it could be either part of a larger dialect ranging over additional provinces, or, conversely, an aggregation of smaller dialects wholly contained within the region. Much more data is required before an empirically adequate answer can be provided to the question whether Riau Indonesian is a geographically coherent entity.

So in this specific sense, then, the reality of Riau Indonesian has yet to have been convincingly demonstrated. But this does not mean that the language provisionally labeled as Riau Indonesian is anything other than a proper language; as argued in this

paper, it is as real as any other language or dialect. All that is missing is a clear picture of where it begins and ends in geographical space. My work on Riau Indonesian is based largely on data collected from a few specific locations: the islands of Bintan, Batam and Karimun, as well as the village of Sungai Pakning and the city of Pekanbaru on the mainland of Sumatra. In lieu of further data from additional places in neighboring provinces, it makes sense to refer to the dialect of Indonesian spoken in these locations as Riau Indonesian, with the understanding that the term is being used provisionally, and that additional much-needed work on varieties of colloquial Indonesian in other parts of Sumatra may possibly lead to a reconsideration of its appropriateness and its replacement with some other term or set of terms. But in this respect, the situation with respect to Riau Indonesian is no different from that which obtains in many other linguistically-understudied places throughout the world, whereby initial exploratory descriptions of a particular speech variety assign it a name, often derived from a toponym, only to modify the name at a later stage when more information on neighboring speech varieties becomes available.

5. OTHER COUNTRIES, OTHER LANGUAGES

Having suggested, in Section 3 above, that from a sociolinguistic point of view, Riau Indonesian is not at all exceptional, but, rather, resembles, in its broad outlines, other varieties of colloquial Indonesian spoken in other regions, one may now ask whether Malay/Indonesian as a whole is exceptional among the languages of the world with respect to its sociolinguistic complexity, or whether other languages also have "their own Riau Indonesians", that is to say, specific varieties defined in terms of the intersection of sociological and geographical properties in ways similar to that of Riau Indonesian. While it does indeed seem clear that the sociolinguistic complexity of Malay/Indonesian is significantly greater than that of many other languages, it is not at all obvious that Malay/Indonesian is that exceptional when compared with other major world languages with similarly large populations of speakers. In fact, there would seem to be good reason to believe that many or most of the world's major languages are associated with comparable degrees of sociolinguistic complexity.

By far the best-documented and best-studied language is English. In addition to the usual descriptions and analyses of Standard English, there is a venerable tradition of studying the classical English dialects, originally centered in Great Britain but subsequently, with the work of Labov and others, successfully transplanted to the USA and elsewhere. Alongside traditional dialectology, there is also a new and burgeoning field devoted to "World Englishes", varieties of English spoken in other parts of the world, in places as diverse as Ghana and the Philippines. And of course, this in turn blends into yet another large and closely related discipline, that of English-lexifier creolistics. Within this plethora of studies of English and English-originated varieties, it is not too hard to come up with plausible if not completely precise analogues of Riau Indonesian. For example, one might venture the suggestion that Standard Indonesian is to Riau Indonesian as Riau Malay is to Standard English as Standard English is to the traditional dialects of London and the surrounding region. Or perhaps as Standard English is to a major urban dialect of Scotland is to Scots. Clearly, these two examples from Great Britain are not parallel to each other; a fortiori, there will be even more differences between them and the Riau case. Nevertheless, it is helpful to think of Riau Indonesian as being *grosso modo* the same kind of beast as Estuary English or an urban

Scottish dialect. In particular, it is likely that if analogues to the 12 claims cited in (1) were proposed with respect to either Estuary English or some major urban dialect of Scotland, most or all would be rejected, quite readily, on grounds more or less similar to those which were invoked, in Section 2 above, in the context of Riau Indonesian.

Similarly complex linguistic landscapes are probably characteristic of many other of the world's major languages. Without specialist knowledge of other regions, or even — in many cases — access to such knowledge, I am not in a position to make sweeping claims. Still, it is likely that many other languages have their own more or less precise analogues of Riau Indonesian, and that in at least some such cases, the varieties in question are underdescribed, or, worse, their very existence unacknowledged. One wonders what kinds of Mandarin are spoken in a city like Kunming, as a more general lingua franca alongside the local variety of Southern Mandarin, or in Guangzhou, where the main local language is Cantonese, or even in Jakarta, where many ethnically Chinese speakers are in the process of switching from languages such as Hokkien, Teochew and Hakka to Mandarin. Or what kinds of Hindi/Urdu are spoken in a place like Jaipur, as a more general koiné alongside the closely related Rajasthani, or Kolkata, where the main local language is Bangla, or London, as a common language shared by part of the south-Asian immigrant community. At least some of these, as well as any number of other similar cases, are quite likely to turn out to be language varieties associated with a population of native speakers, exhibiting stable lexicons and grammars, and used in a wide variety of communicative contexts, or in other words, regular languages and dialects. And it is a safe bet that at least some of the aforementioned language varieties have not yet been described, or had their existence acknowledged, or even been given a name.

If the take-home message from this paper for the Indonesian specialist is that Riau Indonesian exists as a real language variety, alongside many other similar varieties of colloquial Indonesian associated with other parts of the country, the corresponding message for the general linguist is that languages may exhibit greater sociolinguistic complexity than is sometimes assumed, involving a wider range of historical and sociolinguistic types, and that field workers in a particular location should accordingly seek out all of the language varieties that are present there, rather than limiting their attention to just a favoured one. It is a peculiar irony that, as with Riau and other varieties of colloquial Indonesian, it is often the most widespread variety of the language, the dialect that presents itself first to researchers upon arrival, which ends up being unnoticed and accordingly undescribed. As they rush upriver into the jungle to record the speakers of an exotic and possibly endangered language, field workers sometimes pay insufficient attention to the speech of their boatman and his mates.

APPENDIX

This appendix provides supporting data for Tables 1-4 in Section 3 and Figure 1 in Section 4. (In addition to the language abbreviations *Min*, *SkM* and *StI* in Tables 1-4, *RI* is used here as an abbreviation for Riau Indonesian.) Whereas for RI, Min and StI the data is presented in standard orthography, for SkM it is presented in a recently-developed specialized orthography that reflects its richer vowel system; in this orthography, *e* stands for [ə]; *é* for [e]; *è* for [ɛ]; *ó* for [o]; and *ò* for [ɔ].

Table 1: Shared and Distinctive Lexical Features

RI forms cited in the last column are presumed to occur also in Min, SkM and StI if, in those languages, forms exist that are identical in meaning, and similar in form up to the level of automatic and transparent sound correspondences. For example, in row 1 the RI form *makan* is indicated as being present also in SkM even though, in the latter language, the second vowel is fronted, resulting in a form [makæŋ], in accordance with a subphonemic rule specific to SkM (which in fact is alluded to in Table 2 row 3).

Row 1 'eat': RI *makan*; Min *makan*; SkM *makan* (subject to the subphonemic alternation mentioned above); StI *makan*.

Row 2 'look': RI *tengok*; Min *tengok* (also *caliak*); SkM *téngók*; StI *lihat* (in StI *tengok* occurs with a different, albeit related meaning, visit, while *lihat* has a somewhat wider range of uses than RI *tengok*, corresponding also to 'see'.)

Row 3 'yawn': RI *kuap*; Min *kuok* (regular sound correspondence); SkM *sangap*; StI *kuap*.

Row 4 'grue' (the colour term denoting the disjunction of blue and green): RI *hijau*; Min *hijau*; SkM none (*hijau* refers only to 'green', and *biru* to 'blue'); StI none (*hijau* refers only to 'green', and *biru* to 'blue').

Row 5 'fish': RI *ikan*; Min *lauak*; SkM *ikan*; StI *ikan*.

Row 6 'wooden boat with engine': RI *pompong*; Min no specific term; SkM *pómpóng*; StI no specific term.

Row 7 'trousers': RI *celana*; Min *sarawa*; SkM *seluò*; StI *celana*.

Row 8 'give': RI *kasi*; Min *agiah*; SkM *beri*; StI *beri*.

Table 2: Shared and Distinctive Phonological Features

In this table, and in the comments that follow, phonemic representations are enclosed in // and phonetic representations in []. In five of the rows, a parenthetical qualification "in ordinary words" indicates that the phonological feature in question does not apply to words that belong to a number of exceptional classes, such as loan words, words formed by special neologistic processes involving truncation, acronyms and the like, and words of an onomatopoeic or sound-symbolic nature — see Row 5 below for illustration.

Row 1: In all four languages, /c/, /j/ and /ŋ/ occur freely in word-initial and word-medial position, but not in word-final position (where, depending on the language, they are either completely absent, or present only in words belonging to the above-mentioned exceptional classes).

Row 2: In RI, as argued in Gil (2003b, 2006a), there is no lexical stress; similar arguments hold also for Min and SkM. In contrast, StI is generally described as having lexical stress on the penultimate syllable, except when it is a schwa, in which case the stress falls on the final syllable. (In reality, however, the StI facts are probably more complex, with different speakers exhibiting different stress patterns reflecting those of their respective native languages.)

Row 3: In SkM, /a/ is fronted to [æ] before coronal consonants /t/, /s/, /n/ and /l/ in word final position. This does not occur in RI, Min and StI.

Row 4: In RI, there is no phonemic /ə/. Some speakers may have a phonetic [ə] in the penultimate syllable of words such as *ketan* 'sticky rice', but this is in free alternation with [e] and can be argued to be the result of epenthesis applying to an underlying form /ktan/. In Min too there is no phonemic /ə/, though here there is no phonetic [ə] either. In contrast, both SkM and StI clearly have a phonemic /ə/.

Row 5: In RI, SkM and StI, /t/ occurs freely in word-final position. However, in Min, /t/ does not occur word-finally, except in words belonging to the exceptional classes

mentioned in the general comments above, for example *Rahmat*, a proper name borrowed from Arabic; *gatot* 'boast', derived by irregular truncation from *gadang otak* 'big brain'; and *dangdut*, a sound-symbolic name referring to a genre of popular music.

Row 6: In RI, a phonotactic constraint dictates that high vowels may not occur in word-final closed syllables (except as qualified in Row 8 below, or in words belonging to the above-mentioned exceptional classes); this constraint is upheld by a productive phonological rule that lowers the relevant vowel. Thus, for example, the word *patin*, 'silver catfish', underlyingly /patin/, is realized as [patɪn]. The same phonotactic constraint exists also in SkM, though its implementation follows a different path: here the corresponding vowel is phonemically high-mid, as in /paten/. In contrast, in Min and StI, there is no such constraint: high vowels occur freely in final closed syllables, as in /patin/, realized [patin]. (Though, once again, it is necessary to acknowledge that in StI, the facts may vary for speakers with different native accents.)

Row 7: In RI and in StI, /r/ occurs freely in all positions. In contrast, in Min and in SkM, /r/ occurs only word-initially and word-medially; in word-final position it is absent (except in words belonging to one of the exceptional classes).

Row 8: In RI, a rule of vowel harmony overrides the phonotactic constraint referred to in Row 6 above in those cases where the penultimate vowel is identical to the final one; in such instances, lowering does not apply. For example, contrasting with words such as /patin/ realized as [patɪn], are words such as 'candle', /lilin/ realized as [lilin], not [lilɪn]. This rule of vowel harmony does not apply in SkM, where the word for 'candle' is /lilen/, and it is irrelevant for Min and StI, where the constraint against high vowels in word-final position does not apply in the first place.

Table 3: Shared and Distinctive Morphosyntactic Matter Features

As in Table 1, RI forms cited in the last column are presumed to occur also in Min, SkM and StI if, in those languages, forms exist that are identical in meaning, and similar in form up to the level of automatic and transparent sound correspondences.

Row 1: In all four languages, patient-orientation is expressed by the generalized passive marker *di-* (though its precise range of functions may differ from language to language; see Gil 2002b).

Row 2: In RI, Min and SkM, agent-orientation may be expressed by the generalized active marker *N-*, whose morphophonemic realization involves some form of nasal accretion or mutation, depending on the initial segment of the stem to which it attaches. (Note, however, that in Min, there are alternative ways of expressing agent-orientation, as reflected in Table 4 Row 5.) In contrast, in StI, agent-orientation is not expressed by *N-* but rather by the generalized active marker *meN-*.)

Row 3: In RI, Min and Std I, *kalian* occurs as a 2nd person plural pronoun. In SkM, however, the corresponding form is *miko*.

Row 4: In RI and Min, *ndak* is used as a negative marker. While in RI other negative markers, *tak* and *nggak*, are also present, in Min it is the sole marker, and is an abbreviated form of the longer *indak*. In contrast, in SkM and StI, *ndak* is absent, and other forms are used instead, such as *dak* in SkM, *tidak* in StI, and *tak*, shared by SkM and StI, as well as RI. (The profusion of negative markers in Malay/Indonesian is also reflected in 3 out of the 4 items in the morphosyntactic matter column in Figure 1.)

Row 5: In RI, SkM and StI, *nanti* is used to express proximal future time. (Whereas in RI and StI it is the primary form fulfilling this function, in SkM the more common form is

karang and its variants *kang*, *rang* and *ang*). In contrast, in Min, *nanti* is absent, and the corresponding form is *beko*.

Row 6: In RI and SkM, *betul* is used as an intensifier of property words, corresponding to 'very', whereas in Min and StI it is not used with that meaning. However, in all four languages, *betul* has other usages, such as 'real' or 'right'. Also, in all four languages, there are a variety of other forms which share the function of intensifier. While in RI, *sekali* is the most commonly used intensifier, with *betul* as a less frequent variant, in SkM, *betul* is the most frequently occurring intensifier. In Min, *bana* is the most common intensifier, while in StI it is *sangat* or *amat*.

Row 7: In Riau and StI, the content interrogative 'when' is expressed with *kapan*, whereas in Min the corresponding form is *bilo*, and in SkM *biló*.

Row 8: In all four languages, there are one or more applicative suffixes or enclitics expressing increased transitivity, causative and benefactive; however, RI alone of the four makes use of the form *-in* for this function. In actual fact, in RI, *-in* is a less frequent variant of the more commonly used *-kan*, and is associated with more urban, upwardly mobile and trendy language registers. In contrast, SkM has *-kan* as the only applicative marker, Min has either *-kan* in more formal registers, or *-an* more colloquially, and StI has either *-kan* or *-i* depending on a variety of lexical and grammatical factors.

Table 4: Shared and Distinctive Morphosyntactic Pattern Features

Row 1: In all four languages, numerals quantifying discrete individuated objects such as people, animals, houses, etc., may occur in construction with a sortal numeral classifier, but need not necessarily do so, the presence or absence of the classifier being dependent on a variety of semantic and discourse factors; for example RI *tiga (ekor) ikan*, Min *tigo (ikua) lauak*, SkM *tigó (éków) ikan*, StI *tiga (ekor) ikan* for 'three CLF fish', where RI *ekor* and its cognates, literally 'tail', are numeral classifiers used for counting animals.

Row 2: In RI, Min and SkM, the form *do* (in SkM *dó*) is used as an optional negative polarity item, at the end of a phrase containing a negative marker; for example RI *Rinol tak makan ikan (do)*, Min *Rinol indak makan lauak (do)*, SkM *Rinòl tak makan ikan (dó)* literally 'Rinol NEG eat fish NEG.POL', for 'Rinol didn't eat any fish'. StI does not have *do*, or any other marker with such function.

Row 3: In RI, Min and StI, the negation of 'want' is formed in regular fashion, with the negative marker preceding the word for 'want': RI *tak mau*, Min *indak nio*, StI *tidak mau*. In contrast, in SkM, the corresponding collocation * *tak nak* does not occur; its place is taken by the suppletive form *tendak*.

Row 4: In RI and Min, the numeral 'one' may be used as a sentence final particle in the context of a request or polite imperative, for example RI *ikut satu*, Min *ikuik ciek*, literally 'follow one', for 'Can I come with you'. In SkM and StI there is no corresponding usage of the numeral 'one'.

Row 5: In RI, SkM and StI, there is a single generalized active marker expressing agent orientation: in RI and SkM it is *N-*, mentioned in Table 3 Row 2, while in StI it is *meN-*. In contrast, in Min, there are three such markers, *N-*, *ma-*, and *maN-* (though their number may conceivably be reduced to two, if *maN-* is analyzed as a combination of *ma-* and *N-*), the choice between them being dependent on a variety of phonological, grammatical and stylistic factors (Crouch 2009).

Row 6: In RI and SkM, meanings such as 'Kudin and I' are expressed by the collocation of the proper noun with a 1st person exclusive pronoun plus an additional grammatical marker such as the numeral 'two'; since the 1st person exclusive pronoun includes the

proper noun in its reference, the construction is sometimes called an inclusory plural (Haspelmath 2004). For example, in RI 'Kudin and I' is expressed as *kami dua Kudin*, and in SkM as *kami duó Kudin*, literally '1PL:EXCL two Kudin'. (If more than two people are involved, the numeral 'two' is replaced with a general comitative marker, in RI *sama* or *dengan*, in SkM *samó* or *dengan*.) In contrast, in Min and StI, there is no inclusive plural construction, and instead the proper noun occurs in construction with the 1st person singular pronoun: Min *aden jo Kudin* '1SG with Kudin', StI *aku dan Kudin* '1SG and Kudin'.

Row 7: In all four languages there is an enclitic whose general meaning is that of association: attached to a word W, the meaning of W-enclitic can be expressed as 'W associated with x', where x is a variable whose reference is determined by a combination of grammatical, discourse, and extralinguistic context. In RI, the form of the enclitic is *-nya*; attached to a word such as *buku* 'book', the resulting *bukunya* means 'book associated with x', which in turn, depending on context, can be rendered into English as, among others, 'his book', 'our book', 'the library's book', 'the book mentioned in the preceding conversation', 'the book characteristically associated with the given situation', or simply 'the book'. In StI, the form of the associative clitic is also *-nya*, and it has a similar if not quite identical range of usages. However, in both Min and SkM, there are two different forms of the associative clitic, the choice between them depending at least in part on grammatical and discourse factors at present not fully understood: in Min, the two forms are *-nyo* and *-e*, while in SkM they are *-nyó* and *-N*, the latter representing a homorganic nasal consonant.

Row 8: In RI there is a construction of the form *X kek Y kek ...* with the interpretation 'X, Y ... or associated things', for example, *buku kek koran kek* 'books, newspapers, or things like that'. (In this construction, any number of items can be conjoined.) An appropriate name for this construction, which I have not seen mention of in the general linguistic literature, might therefore be associative disjunction. However, none of the other three languages have a similar construction, with *kek* or any other marker.

Figure 1: Some Malay/Indonesian Isoglosses between Kuala Lumpur and Jakarta

The dialects associated with the locations represented in Figure 1 are referred to below with the following abbreviations: *KLM* Kuala Lumpur Malay, *PI* Palembang Indonesian, *KI* Kalianda Indonesian, *JI* Jakarta Indonesian.

Phonological Features:

Isogloss 1: To the north, word-final /r/ is absent; to the south it is present. For example: KLM /paga/, RI, PI, KI, JI /pagar/. (This is the same feature referred to in Table 2 row 7.)

Isogloss 2: To the north, word-final /ə/ in closed syllables is absent; to the south it is present. For example: KLM, RI /malam/; PI, KI, JI /maləm/ 'night'. (In KLM, this constraint applies only to native words; in loan words, /ə/ may occur in a final closed syllable, for example /hɛnsəm/ 'handsome'.)

Isogloss 3: To the north, word final /k/ is realized as [ʔ]; to the south as [k]. For example, KLM, RI, PI [masaʔ]; KI, JI [masak] 'cook'. The presence of an underlying /k/ in KLM, RI and PI is evidenced by alternations such as [masaʔ] ~ [masakan] 'cuisine'. (Again, in KLM, this constraint applies only to native words; in loan words, [k] may occur word-finally, for example KLM [cɛk] '(bank) check', RI, PI [ojek] 'motorcycle taxi', from Javanese, probably via JI.)

Isogloss 4: To the north, word-final /a/ is only ever realized as a central vowel ([ɨ̜] in KLM, [a] in RI, PI and KI); to the south, it has a stylistically-marked alternative realization as mid-front [e]. For example, from 'what' /apa/, KLM [apɨ̜], RI, PI, KI [apa], JI [apa] ~ [ape].

Morphosyntactic Matter Features:

Isogloss 1: To the north, the negative marker *nggak* is absent; to the south it is present.

Isogloss 2: To the north, the distal demonstrative *no* is absent; to the south it is present. (Deictics in Malay/Indonesian generally occur in (at least) the following three series: simple demonstrative, complex demonstrative and complex locative, and typically express a three-way distinction between proximal, medial and distal, eg. in JI, simple demonstratives *ni, tu, no*; complex demonstratives *ini, itu, ono*; and complex locatives *sini, situ, sono*. However, in many dialects, the paradigm is defective, with simple and complex demonstratives lacking a distal, eg. in RI, simple demonstratives *ni, tu*, [none]; complex demonstratives *ini, itu*, [none]; complex locatives *sini, situ, sana*. Historically, the defective paradigm is apparently the original one, with the JI forms created by analogy and under Javanese influence and the simple distal demonstrative *no* subsequently spreading north from JI to KI and PI via language contact.)

Isogloss 3: To the north, the negative marker *tak* is present; to the south it is absent.

Isogloss 4: To the north, the negative marker *kagak* is absent; to the south it is present.

ACKNOWLEDGEMENTS

This paper owes its existence to many friends and colleagues who have challenged my work on Riau Indonesian in various ways. If I have chosen to avoid attributing certain claims to specific persons, this is in order to emphasize that, rather than trying to score points off each other, we share the common goal of advancing our understanding of the issues under discussion. Indeed, my own understanding of what Riau Indonesian is has profited immensely from having to defend my case against the claims discussed in this paper. The main arguments in this paper were first put forward at the Seventh International Symposium on Malay/Indonesian Linguistics, Nijmegen, The Netherlands, 28 June 2003; I am grateful to members of the audience there for their helpful comments.

REFERENCES

- Cole, Peter and Gabriella Hermon (2005) "Subject and Non-Subject Relativization in Indonesian", *Journal of East Asian Linguistics* 14:59-88.
- Cole, Peter, Gabriella Hermon and Yanti (2008) "Voice in Malay/Indonesian", *Lingua* 118/10:1500-1553.
- Collins, James T. (1980) *Ambonese Malay and Creolization Theory*, Dewan Bahasa dan Pustaka, Kuala Lumpur.
- Crouch, Sophie (2009) *The Discourse and Pragmatic Effects on Voice and Verbal Morphology in Minangkabau*, MA thesis, University of Western Australia, Perth.
- Donohue, Mark (to appear) "Papuan Malay".
- Donohue, Mark and Tim Denham (to appear) "Island Southeast Asia During the Mid-Holocene: Reframing Austronesian History", *Current Anthropology*.
- Donohue, Mark and Yusuf Sawaki (2007) "Papuan Malay Pronominals: Forms and Functions", *Oceanic Linguistics* 46:253-276.

- Durvasula, Karthik (2008) "Multiple Categorical Sources for Surface Partially-Nasal Stops and the Nature of their Variability", Paper presented at the 31st Annual Colloquium of Generative Linguistics in the Old World (GLOW), Toulouse, France, 25 March 2008.
- Dryer, Matthew (2005) "Order of Genitive and Noun" in M. Haspelmath, M. Dryer, D. Gil and B. Comrie eds., *The World Atlas of Language Structures*, Oxford University Press, Oxford, 350-353.
- Gil, David (1994) "The Structure of Riau Indonesian", *Nordic Journal of Linguistics* 17:179-200.
- Gil, David (1999) "Riau Indonesian as a Pivotless Language", in E.V. Raxilina and Y.G. Testelec eds., *Tipologija i Teorija Jazyka, Ot Opisanija k Objasneniju, K 60-Letiju Aleksandra Evgen'evicha Kibrika* (Typology and Linguistic Theory, From Description to Explanation, For the 60th Birthday of Aleksandr E. Kibrik), Jazyki Russkoj Kul'tury, Moscow, 187-211.
- Gil, David (2000) "Syntactic Categories, Cross-Linguistic Variation and Universal Grammar", in P. M. Vogel and B. Comrie eds., *Approaches to the Typology of Word Classes*, Empirical Approaches to Language Typology, Mouton, Berlin and New York, 173-216.
- Gil, David (2001a) "Creoles, Complexity and Riau Indonesian", *Linguistic Typology* 5:325-371.
- Gil, David (2001b) "Escaping Eurocentrism: Fieldwork as a Process of Unlearning", in P. Newman and M. Ratliff eds., *Linguistic Fieldwork*, Cambridge University Press, Cambridge, 102-132.
- Gil, David (2002a) "Ludlings in Malayic Languages: An Introduction", in Bambang Kaswanti Purwo ed., *PELBBA 15, Pertemuan Linguistik Pusat Kajian Bahasa dan Budaya Atma Jaya: Kelima Belas*, Unika Atma Jaya, Jakarta, 125-180.
- Gil, David (2002b) "The Prefixes di- and N- in Malay / Indonesian Dialects", in F. Wouk and M. Ross eds., *The History and Typology of Western Austronesian Voice Systems*, Pacific Linguistics, Canberra, 241-283.
- Gil, David (2003a) "Colloquial Indonesian Dialects: How Real Are They?", Paper presented at the Seventh International Symposium on Malay/Indonesian Linguistics, Association for Linguistic Typology, Nijmegen, The Netherlands, 28 June 2003.
- Gil, David (2003b) "Intonation Does Not Differentiate Thematic Roles in Riau Indonesian", in A. Riehl and T. Savella eds., *Proceedings of the Ninth Annual Meeting of the Austronesian Formal Linguistics Association (AFLA9)*, Cornell Working Papers in Linguistics 19:64-78.
- Gil, David (2004a) "Learning About Language from Your Handphone; *dan*, *and* and *&* in SMSs from the Siak River Basin", in Katharina Endriati Sukatmo ed., *Kolita 2, Konferensi Linguistik Tahunan Atma Jaya*, Pusat Kajian Bahasa dan Budaya, Unika Atma Jaya, Jakarta, 57-61.
- Gil, David (2004b) "Riau Indonesian *sama*, Explorations in Macrofunctionality", in M. Haspelmath ed., *Coordinating Constructions*, Typological Studies in Language 58, John Benjamins, Amsterdam, 371-424.
- Gil, David (2005a) "Can You Drink a Book in Malay/Indonesian? Variation Across Isolect, Socio-Economic Status and Age", Paper presented at the Ninth International Symposium on Malay/Indonesian Linguistics, Association for Linguistic Typology, Ambun Pagi, Indonesia, 27 July 2005.

- Gil, David (2005b) "Distributive Numerals", in M. Haspelmath, M. Dryer, D. Gil and B. Comrie eds., *The World Atlas of Language Structures*, Oxford University Press, Oxford, 222-225.
- Gil, David (2005c) "From Repetition to Reduplication in Riau Indonesian", in B. Hurch ed. *Studies on Reduplication*, Empirical Approaches to Language Typology 28, Mouton de Gruyter, Berlin, 31-64.
- Gil, David (2005d) "Isolating-Monocategorical-Associational Language", in H. Cohen and C. Lefebvre eds., *Categorization in Cognitive Science*, Elsevier, Oxford, 347-379.
- Gil, David (2005e) "Word Order Without Syntactic Categories: How Riau Indonesian Does It", in A. Carnie, H. Harley and S.A. Dooley eds., *Verb First: On the Syntax of Verb-Initial Languages*, John Benjamins, Amsterdam, 243-263.
- Gil, David (2006a) "Intonation and Thematic Roles in Riau Indonesian", in C.M. Lee, M. Gordon, and D. Büring eds., *Topic and Focus, Cross-Linguistic Perspectives on Meaning and Intonation*, Studies in Linguistics and Philosophy 82, Springer, Dordrecht, 41-68.
- Gil, David (2006b) "The Acquisition of Voice Morphology in Jakarta Indonesian", in N. Gagarina and I. Güllow eds., *The Acquisition of Verbs and Their Grammar: The Effect of Particular Languages*, Springer, Dordrecht, 201-227.
- Gil, David (2007) "Creoles, Complexity and Associational Semantics", in U. Ansaldi and S.J. Matthews eds., *Deconstructing Creole: New Horizons in Language Creation*, John Benjamins, Amsterdam, 67-108.
- Gil, David (2008a) "How Complex Are Isolating Languages?" in F. Karlsson, M. Miestamo and K. Sinnemäki eds., *Language Complexity: Typology, Contact, Change*, John Benjamins, Amsterdam, 109-131.
- Gil, David (2008b) "The Acquisition of Syntactic Categories in Jakarta Indonesian", *Studies in Language* 32:637-669.
- Gil, David (2009) "How Much Grammar Does It Take to Sail a Boat", in G. Sampson, D. Gil and P. Trudgill eds., *Language Complexity as an Evolving Variable*, Oxford University Press, Oxford.
- Gil, David (to appear) "Riau Indonesian: A Language without Nouns and Verbs", in J. Rijkhoff and E. van Lier eds., *Flexible Word Classes*, Oxford University Press, Oxford.
- Haspelmath, Martin (2004) "Coordinating Constructions, An Overview", in M. Haspelmath ed. *Coordinating Constructions*, Typological Studies in Language 58, John Benjamins, Amsterdam, 3-39.
- McWhorter, John (2001) "What People Ask David Gil and Why", *Linguistic Typology* 5:388-412.
- McWhorter, John (2005) *Defining Creole*, Oxford University Press, Oxford and New York.
- McWhorter, John (2006) *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*, Oxford University Press, New York.
- McWhorter, John (2008) "Why Does a Language Undress? Strange Cases in Indonesia", in F. Karlsson, M. Miestamo and K. Sinnemäki eds., *Language Complexity: Typology, Contact, Change*, John Benjamins, Amsterdam.
- Minde, Don van (1997) *Malayu Ambong: Phonology, Morphology, Syntax*, CNWS, Leiden.

- Paauw, Scott (2008) *The Malay Contact Varieties of Eastern Indonesia: A Typological Comparison*, PhD dissertation, State University of New York at Buffalo.
- Pakir, Anne Geok-in Sim (1986) *A Linguistic Investigation of Baba Malay*, PhD Dissertation, University of Hawai'i, Manoa.
- Prentice, D.J. (1994) "Manado Malay: Product and Agent of Language Change", in T. Dutton and D.T. Tryon eds., *Language Contact and Change in the Austronesian World*, Mouton de Gruyter, Berlin, 411-441.
- Pullum, Geoffrey K. (1989) "Topic ... Comment: The Great Eskimo Vocabulary Hoax", *Natural Language and Linguistic Theory* 7:275-281.
- Rheeden, Hadewych A. van (1999) "The Emergence and Structure of Steurtjestaal, A Mixed Orphanage Language of Colonial Java", Paper presented at the Third International Symposium on Malay/Indonesian Linguistics, Association for Linguistic Typology, Amsterdam, The Netherlands, 25 August 1999.
- Slomanson, Peter (to appear) "Substrate Features in Sri Lankan Malay", in C. Lefebvre ed., *Creoles and Their Substrates* [tentative title], John Benjamins, Amsterdam.
- Sneddon, James Neil (2006) *Colloquial Jakartan Indonesian*, Pacific Linguistics, Canberra.
- Steinhauer, Hein (1983) "Notes on the Malay of Kupang (Timor)", in J.T. Collins ed., *Studies in Malay Dialects, Part II, NUSA 17*, Universitas Atma Jaya, Jakarta, 42-64.
- Stoel, Ruben (2005) *Focus in Manado Malay: Grammar, Particles and Intonation*, CNWS Publications, Leiden.
- Tadmor, Uri (2005) "What Kind of Linguistic Creature is Semarangan?", Paper presented at the Ninth International Symposium on Malay/Indonesian Linguistics, Association for Linguistic Typology, Ambun Pagi, Indonesia, 28 July 2005.
- Tadmor, Uri (2007) "Local Elements in Betawi Malay", Paper presented at the First International Symposium on the Languages of Java, Semarang, Indonesia, 16 August 2007.
- Taylor, Paul M. (1983) "North Moluccan Malay: Notes on a 'Substandard' Dialect of Indonesian", in J.T. Collins ed., *Studies in Malay Dialects, Part II, NUSA 17*, Universitas Atma Jaya, Jakarta, 14-27.
- Tjung, Yassir Nasanius (2006) *The Formation of Relative Clauses in Jakarta Indonesian, A Subject-Object Asymmetry*, PhD dissertation, University of Delaware, Newark.
- Voorhoeve, C. L. (1983) "Some Observations on North-Moluccan Malay", in J.T. Collins ed., *Studies in Malay Dialects, Part II, NUSA 17*, Universitas Atma Jaya, Jakarta, 1-13.
- Wouk, Fay (1989) *The Impact of Discourse on Grammar: Verb Morphology in Spoken Jakarta Indonesian*, PhD dissertation, UCLA, Los Angeles.
- Wouk, Fay (1999) "Dialect Contact and Koineization in Jakarta, Indonesia", *Language Sciences* 21:61-86.
- Yanti (in preparation) *Jambi Malay*, PhD Dissertation, University of Delaware, Newark.

ABBREVIATIONS

This paper makes use of the following glossing abbreviations: CLF classifier, EXCL exclusive, NEG negative, PL plural, POL polarity, SG singular, 1 first person, 2 second person.