1 Introduction
The Pacoh pronoun system is exceptional among the Mon-Khmer languages in Southeast Asia for having morphologically distinct case-marked sets, including subjects/direct objects, indirect objects, and possessive pronouns (Watson 1964, Alves 2000). Moreover, these pronoun sets have gone further along the path of a grammatical cline, serving a number of semantico-syntactic functions beyond their capacity as pronouns. It is the purpose of this study to show these paths of grammaticalization and how they represent a number of linguistic processes in the syntactic evolution of lexical material. In particular, while semantic changes are present, the changes more often follow step by step along parameters of both semantic fields and syntactic features.

The sections of this study include (1) general discussion of grammaticalization with additional information on relevant aspects of Southeast Asian lexical subcategories, (2) a summary of the basic Pacoh pronoun system with overt case-marking affixes, (3) a section on Pacoh pronouns grammaticalized as general possessive or dative relator nouns, (4) discussion of the Pacoh pronoun grammaticalized as a marker of plurality, (5) discussion of Pacoh pronouns grammaticalized as conjunctions, and (6) a brief look at related derivational patterns in other Mon-Khmer languages. In each section, both the semantic and syntactic aspects of each derivational relationship are discussed where those aspects are seen as distinct. A constant issue among all the grammatical vocabulary derived from Pacoh pronouns is whether or not there are semantic restrictions on co-occurring nouns, namely whether or not those nouns must be human. In addition, some of the aspects of the cognitive characteristics of the changes are shown in relation to the types of universal changes seen in other languages.

2 Grammaticalization and Syntactic Issues
The term ‘grammaticalization’ refers ultimately to a subcategory of (a more general process of) semantic shift in which words or word parts become increasingly abstract and serve grammatical functions (Hopper and Traugott 1993, Heine and Kuteva 2002). In fact, the term ‘grammatical’ itself does not have a clear breaking point in the literature on the topic, and this definition thus becomes circular (i.e., Grammaticalization is the process of becoming grammatical. Grammatical words undergo grammaticalization). Still, intuitively, identifying function words/grammatical vocabulary appears to be a less controversial matter.

In the end, perhaps the best way to deal with this issue is to acknowledge that the definition is somewhat imperfect and deal with the rough edges later. However, to attempt to clarify this issue somewhat, distinctive features (both semantic and syntactic) can be used to bring the issue of ‘grammatical’ and hence ‘grammaticalization’ into a sharper fo-
The purpose of the following subsections is to clarify the kinds of features used in this paper, the kinds of clines seen among Pacoh pronouns, and the key concepts of relator nouns, a class of words into which the pronouns have grammaticalized.

2.1 Degree of Grammaticality and the Use of Syntactic Features

While the differences between the extremes of function and content words seem clear, there does appear to be a degree of grammaticality in which there must be areas of uncertainty. While the word ‘dog’ (referring to the animal) intuitively serves non-grammatical content word, and words or word parts expressing ‘past tense’ are intuitively grammatical in nature, some words and concepts appear to be somewhere in between. Examples are time words (such as ‘today,’ which is a true time noun with specialized distribution, as opposed to ‘day,’ which is a common countable noun that requires prepositions to mark their function in time clauses, such as ‘on that day’ or ‘during those days’) and causative words (such as ‘make’ which differs from ‘force’ as the latter requires a ‘to’ infinitive while the former requires only a bare non-finite verb). No specific line has been drawn between grammatical and non-grammatical elements, and this is perhaps due to the semantic definitions typically given for grammaticalization. Typically, in the literature on grammaticalization, semantico-syntactic categories are used, with somewhat more emphasis on the cognitive and semantic side. However, using semantic fields is not the only way to identify grammaticality and hence grammaticalization. A concrete partial solution lies in the use of syntactically-grounded lexico-syntactic features, features that manifest themselves in freedom or restriction in syntactic distribution.

Parts of speech represent distinct grammatical categories and can be rated according to degree of grammaticalization to identify chains of grammaticalization and account for the direction of the shifts (though not necessarily the cause(s) of the shifts). Consider the differences between common nouns, pronouns, and classifier nouns. Common nouns in Pacoh (and other Southeast Asian languages) have the fewest grammatical restrictions, occurring as subjects, objects of verbs, objects of prepositions, possessive attributes, or semantic heads of quantified noun phrases. Pacoh pronouns have roughly the same functions, though they have anaphoric reference and special morphological attributes which limit restrict their distributions and semantico-syntactic functions. Classifiers in Pacoh, which have even more specialized distributional properties, are a subcategory of nouns that are often derived from nouns and can take demonstrative complements, a characteristic common to nouns.

<table>
<thead>
<tr>
<th>Common Nouns</th>
<th>Pronouns</th>
<th>Classifier Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>nouns</td>
<td>nouns</td>
<td>nouns</td>
</tr>
<tr>
<td>fewest restrictions</td>
<td>some restrictions</td>
<td>most restrictions</td>
</tr>
<tr>
<td>open set</td>
<td>closed set</td>
<td>closed set</td>
</tr>
</tbody>
</table>

Still, within these primary lexical categories, there are additional degrees of grammaticality. Increasing restrictions and specialization can be rated with polar features, as in Table 1. We can see on the one end common nouns as the least marked and least ‘gram-
mational’ and the classifier nouns\(^1\) as the most marked and thus most ‘grammatical’. The features themselves correspond well to concepts of markedness.

Parts of speech themselves can be roughly graded, verbs, nouns, and adjectives being less grammatical than prepositions, conjunctions, articles, and sentence particles. Indeed, in many cases, words from the former group are the derivational source for those in the latter. Thus, we can get a general sense of what constitutes grammatical vocabulary. Markedness in terms of semantic features, syntactic constraints, as well as the number of items belonging to a lexical class or subclass all give us some idea of what words are more grammatical than others.

While acknowledging some circularity in the logic, the fact that grammatical functions (indicating specific kinds of relationships between words, phrases, or even larger units of speech) can be clearly identified and explicitly referred to demonstrate that these linguistic changes can also be differentiated from non-grammatical semantic changes, say for example, the relationship between ‘dog’ as a noun and as a verb, an example of general semantic extension.\(^2\)

### 2.2 Summary of Clines in Pacoh

Four kinds of grammaticalization of Pacoh pronouns are considered in this study, as summarized in Table 2, which contains categories and formalism used in Heine and Kuteva’s (2002) catalogue of grammaticalization chains. With the exception of Table 2, this article does not employ the same categories and directions of grammaticalization used by the work of Heine and Kuteva, in which all capital letters and arrow symbols indicate the direction of change. Instead, syntactic properties are indicated loosely with commonly used grammatical terms, such as ‘noun’ and ‘preposition’, ‘possessive’ and ‘dative’, and ‘human’ and ‘plural’.

### Table 2: List of Grammaticalization Clines of Pacoh Pronouns

<table>
<thead>
<tr>
<th></th>
<th>Pronoun</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>?a.daria</td>
<td>3rd DATIVE &gt; DATIVE</td>
</tr>
<tr>
<td>2</td>
<td>?a.daria</td>
<td>3rd POSSESSIVE &gt; POSSESSIVE</td>
</tr>
<tr>
<td>3</td>
<td>?a.pe</td>
<td>3rd PLURAL &gt; PLURAL</td>
</tr>
<tr>
<td>4</td>
<td>various</td>
<td>PRONOUN &gt; CONJUNCTION</td>
</tr>
</tbody>
</table>

While Heine and Kuteva provide an extensive list of types of grammaticalization among languages of the world, Pacoh (and other Southeast Asian languages, as discussed in § 4) show a few clines not mentioned by those authors, though the Pacoh clines discussed here use basic categories seen as common sources for grammaticalized forms, mostly 3rd person personal pronouns. Of the four categories discussed in this paper, only category 3 (PLURAL) is listed in Heine and Kuteva, although category 4 is similar to their category DEMONSTRATIVE > CONJUNCTION since 3rd person pronouns are semantic correlates of

---

1 This author considers what are commonly called classifiers, measure words, and the like to be a subclass of nouns. Justification for this is in Sak-Humphrey 1996 and Alves 2000.

2 Though a type of grammatical change nonetheless that is paralleled in other parts of English in which nouns are derived as verbs, a kind of N→V rule.
distal demonstratives. Category 1 has some justification as an areal or possibly genetically unmarked type of change within the Mon-Khmer language group as it is seen in another language in another branch of Mon-Khmer, namely, Palaung/Bulang, which is spoken in areas of northern Burma and bordering areas of Yunnan province in China and thus could not have had language contact (see § 4).

What Heine and Kuteva’s formalism, which uses very general semantico-syntactic categories, does not show is that the first three categories preserve the primary part of speech (i.e., noun) and that the last category, while changing part of speech, does preserve features related to the human pronouns, namely that human nouns co-occur with those derived conjunctions, in some cases even requiring a kind of agreement in plurality (see § 3.3). In each case, some properties change, while others are kept. When a more distinct change has occurred, typically, there are intervening steps, changes of single syntactic features which lead to subsequent changes further from source meanings. While these changes can be seen in terms of semantic extension due to implied semantic relationships and reanalysis, the changes in each case follow pre-established syntactic categories and features.

2.3 Relator Nouns in Mon-Khmer Languages and Homophony

Relator nouns, like adpositions, indicate spatial, directional, or possessive relationships between nouns, or they may indicate directions or locations of verbs. Thus, they tend to have semantic properties similar to prepositions in English, but with somewhat different distributional properties. Relator nouns are inalienably possessed (Indrambarya 1984, Sak-Humphrey 1996) and so cannot be connected to morphemes that already carry possessive semantic features (such as Pacoh ūon mentioned in §3.1.1), though they are, as nouns, able to take following demonstratives. Furthermore, they are entirely uncountable and cannot be the semantic heads in quantified noun phrases. Hence, such words serve very specific grammatical purposes and thus have specialized syntactic features and relatively abstract semantic properties.

Pacoh, like most Mon-Khmer languages and other Southeast Asian languages, has a set of locational relator nouns to indicate various substantive locations, such as ‘front’ (as opposed to ‘before’), ‘back’ (as opposed to ‘behind’), and the like. Such words themselves have often down the path of grammaticalization from the semantically concrete to the more abstract (e.g., ‘face’ to ‘front’ to ‘before (time)’), from a specific, concrete part of something to a general area.

The question then is how the source and target meanings in a grammaticalization chain are related. I take the view that the mental lexicon distinguishes between homophonous forms through syntactic and/or semantic distinctions, rather than an approach involving polysemy. Viewing words variously as polysemous or homophonous (e.g., ‘to’ as a movement direction is a different word than ‘to’ in an infinitive but polysemous with ‘to’ as a dative goal) results in ad hoc explanations that cannot be decisively proven or refuted.

---

3 The term ‘relator noun’ has been used largely in literature using the Lexicase dependency theory. Justification for these words as nouns is their ability to take following demonstrative noun dependents and for their functions as locative objects of verbs and prepositions.

4 Pacoh does have a limited set of prepositions. See Alves 2000 for examples.