Title
Correlating Complexity: A Typological Approach

Journal
UC Berkeley PhonLab Annual Report, 1(1)

ISSN
2768-5047

Author
Shosted, Ryan K

Publication Date
2005

DOI
10.5070/P719n0c4m4

Correlating complexity: a typological approach

Ryan K. Shosted
University of California, Berkeley

"In sum, linguists don't even think of trying to rate languages as good or bad, simple or complex."
(O'Grady et al. 2005: 7)

Abstract
Proceeding from the assumption that all languages are equally complex, there exists a corollary, widely held but poorly documented, herein referred to as the negative correlation hypothesis. It states: if one component of language is simplified then another must be elaborated. Here, this assumption is reformulated in terms of a scientific hypothesis and subjected to statistical analysis. Thirty-two geographically diverse languages representing 30 language families and two isolates are tested for syllable count and inflectional synthesis on the verb as a means of rating their phonological and morphological complexity, respectively. The correlation between these measures is found to be slightly positive ($r=0.0704$), but statistically insignificant ($p>0.05$), indicating that the negative correlation hypothesis, if it is to be retained, still awaits scientific confirmation.

1.0. Introduction

1.1. The complexity problem and holistic typology

Among some linguists, it seems there is a tacit assumption that the various components of human language (phonology, morphology, the lexicon, syntax and semantics) correlate negatively within a single language based on complexity. From such an assumption, it could be argued that if a language is simple in terms of its phonology (e.g. if it has a low number of contrastive elements, distinctive features, allophonic alternations, etc.) the language as a system will somehow force elaboration in the areas of syntax or morphology, etc. Throughout this paper, I will refer to this notion as the "negative correlation hypothesis." Introductory linguistics textbooks often note in their first few pages that all grammars are, in a sense, "created equal." For example, Akmajian et al. (1997: 8) declare, "Although it is obvious that specific languages differ from each other on the surface, if we look closer we find that human languages are surprisingly similar. For instance, all known languages are at a similar level of complexity and detail—there is no such thing as a primitive language." In their text, the first stated goal of Cipollone et al. (1998: 2) is to convince the reader that "[e]very language is enormously complex." On the matter of what they call grammatical "equality," O'Grady et al. (1997: 6) reason that "[i]t is futile to attempt to 'rank' languages in terms of relative sophistication."

Nonetheless, it seems clear to even the casual observer that some languages do in fact have smaller/simpler/less sophisticated linguistic components than corresponding components in other languages. For example, Ju|'hoan (Northern Khoisan) has 93 phonemic consonants and Yimas (Lower Sepik) has only twelve (see Table 2). How, then, can these two languages end up with "a similar level of complexity and detail" when all other linguistic components are taken into account? Based on the foundational assumption of "grammatical equality," there are two possible answers: (1) Yimas compensates for what it lacks in terms of phonological complexity through elaboration (beyond that of Ju|'hoan) in terms of its syntax, morphology, semantics, etc.; or (2) There is some aspect of phonological complexity that is not effectively captured by simply counting the consonants in these two languages (i.e. the phonologies are equally complex when other variables are considered). Both are valid responses to the perplexing position we inherit from the dictum that all languages are equally complex (and the corollary, in the case of (1), that their components may not be). However, neither response, to my knowledge, has ever been validated using quantitative methods that produce falsifiable results—an astonishing lacuna, considering the widespread acceptance of grammatical equality.

Several enlightening discussions of differential linguistic complexity have recently come forth. Dahl (2004: 42) observes: "Given that a language as a system can be seen as involving both resources [the inventory of structured linguistic units] and regulations [roughly, 'rules'], it follows that a language could be characterized as more or less complex with respect to both these notions." Kusters (2003: 5) provides an excellent historical overview of the "thoughtless repetition" of the claim that "there are no differences in overall complexity." In the present study, I will attempt to characterize languages according to the complexity of their resources, in the sense of Dahl (2004). I do so in the spirit of combating the "thoughtless repetition" of a claim that has been, until fairly recently, more a matter of dogma than of science.

Antoine Meillet ([1906] 1921: 16) famously opined that languages are *systèmes où tout se tient*. The claim has documented origins older than Meillet's writings and has long tempted students of the mind eager to substantiate relationships between the phonological and morphological components of language. As Plank (1998) discusses, numerous claims have been advanced in the Western linguistic tradition concerning how phonemic inventories, syllable shapes, words and morphemes, phonological rules, tones, accents, and prosodic patterns are systematically interrelated with morphological and syntactic variables. The latter have traditionally included synthetic versus polysynthetic grammar, separatist versus cumulative exponence, complexity of grammatical units, and morpheme, word, and phrase order (Plank 1998: 196). Nonetheless, none of these claims concerning inter-level correlations have found their way into the canon of typological implications (e.g. Greenberg et al. 1978). Concerning the matter of what he calls "holistic typology," Plank (1998: 223) observes that "over the last 250 or so years it has not been established that phonology covaries with morphology and syntax. But then, it has not been established either that it does not." In other words, the research hypothesis of correlation has neither been accepted nor rejected after all this time. One may reasonably wonder how the negative correlation hypothesis persists when it has never, in fact, been proven.

I propose to measure complexity using a quantitative method that produces falsifiable results, determine the correlation between different kinds of complexity, and finally subject the results to standard statistical procedures that assess their significance. Indicators of complexity in the present study are exclusively phonological and morphological. I have consciously left out syntax, pragmatics, the lexicon, and semantics even though, in principle, the negative correlation hypothesis applies to relationships between any two or more of these components as well. There are various reasons for this choice. First, morphological and phonological indicators of complexity can be surveyed in a very straightforward manner. As explained in greater detail below, the surveyor's job in this study is to count structural units such as phonemes, tones, and inflectional markings in a given language. Afterwards, the combinatory possibilities of these units (e.g. in syllables or verbs) are evaluated in order to come up with complexity scores for phonology and morphology. An analogous procedure in syntax and pragmatics, though not impossible to imagine, would face greater obstacles in implementation. This is due to the paucity of diverse typological data outside the domains of morphology and phonology as well as greater disagreement as to what units ought to be counted in the first place. There is no reason that such studies cannot be approached seriously, but a comparison of phonology and morphology seems like a suitable place to begin.

1.2. What is complexity?

As Kusters & Muysken (2001) point out, arguments about complexity are anything but simple. According to Wurzel (2001: 377–378), "There is no single definition of grammatical complexity that is generally accepted; there are several, more or less pre-theoretical notions of complexity, and it is not obvious which one of them is the most adequate, if indeed there is a single most adequate one." Kusters (2003: 6) makes significant strides by defining complexity as the amount of effort a second language learner or "outsider" has to make in order to "become acquainted with the language in question." Indeed, it seems reasonable to take the perspective of the language learner when considering complexity. After all, one is likely to consider a language's complexity as compared to the perceived complexity of one's native tongue. The author goes on to relate "outsider complexity" in the verbal inflection systems of various languages to sociohistorical factors.[1]

Two provocative articles by John McWhorter (2001; 2001a; see also McWhorter 2005) on the matter of grammatical complexity are indeed illuminating. In criticizing the received wisdom that all languages are equally complex (albeit in different modules of linguistic expression), McWhorter argues

[1] In the present study, I emphasize viable methods for sampling the structural units of language and then submitting the results to statistical analysis. Though I commend Kusters' (2003) definition, I will not explore it further by discussing the social and cultural histories of the 32 languages in my sample.

that creoles are the world's simplest grammars. First, he reasons that there are some elements of grammar that are not strictly essential to communication, based on the fact that they are not universally present in human language. These nonessential elements, which creoles do not typically possess to any significant degree, are reasoned to be non-functional or redundant. McWhorter (2001a: 389) actually compares them to human bodily hair, "a matter of habit, doing no harm and thus carried along" for the evolutionary ride. He continues, "Grammars retain both useful and useless features, and meanwhile lose both useful and useless features" (ibid.).

According to McWhorter, languages born in the crucible of extreme and sometimes traumatic sociolinguistic contact—such as pidgins—are not burdened by the accretion of many of the linguistic elements found in older languages. Thus, in McWhorter's characterization, complexity amounts to little more than "overspecification." In this view, languages gradually accrete strategies and "ornaments" of expression that serve no greater, universal need in human communication, e.g. grammatical gender, which is by no means as prevalent in the languages of the world as, for instance, coronal stops. This process eventually results in languages that are encrusted by the weight of much grammatical complexity which happens to be absent in younger creole languages.

McWhorter (2001: 135–137) proposes the design of several metrics for complexity, operative in four modules[2] of linguistic expression (these appear to be roughly isomorphic with the classical subfields of linguistics: phonetics/phonology, syntax, semantics/pragmatics, and morphology). It is McWhorter's "guiding intuition" that "an area of grammar is more complex than the same area in another grammar to the extent that it encompasses more overt distinctions and/or rules than another grammar" (135).

McWhorter's diagnostic for complexity was harshly criticized by a number of detractors (e.g. Hagège 2001; LeFebvre 2001; DeGraff 2001) though supported by Seuren (2001) and Trudgill (2001). The criticisms in particular, along with McWhorter's (2001a) rejoinder, constitute a rich and varied theoretical backdrop for discussing the metric of complexity that will be used in the current investigation. Despite various arguments to the contrary (to be discussed below), it is my position that phonological and morphological complexity can be sampled in reasonable ways. To show the strengths and weaknesses of this position, I will present in the following section a discussion of McWhorter's proposed metrics for morphology and phonology.

1.2.1. Phonological complexity

McWhorter's diagnostic dictates that a greater number of marked phonemes equates to a more complex phonological system. He rejects the notion of markedness in the sense of Trubetzkoy (1931), where markedness is considered a measure of articulatory complexity. Noting that "the precise measurement of articulatory complexity has proven too controversial" for his purposes, McWhorter (2001: 135) opts for a definition of markedness based on cross-linguistic distribution. Thus, for him, marked phonemes are less commonly attested in the languages of the world than other phonemes "conventionally deemed unmarked" (ibid.). It follows that a Khoisan language, with a substantial inventory of rarely attested velaric ingressive consonants should be considered more complex than Yimas, which has only 12 consonants, the least common of which is probably the palatal lateral (Foley 1991). McWhorter would reason that Khoisan phonology is more complex not only because it contains the marked velaric ingressive sounds but because it contains these marked sounds in addition to a great many unmarked ones. In other words, the presence of marked phonemes implies the presence of unmarked counterparts. He considers this implication valid since there are no languages whose phonological inventories contain only cross-linguistically marked phonemes (e.g. a language in which the only consonants are ejective).

Tonal distinctions also factor into McWhorter's phonological metric. He claims that tonal systems become increasingly complex as a function of the total number of their tonal distinctions. Multiple tones and sandhi rules require the speaker-listener to control both articulation and perception of a larger set of contrasts as well as filter out the potentially confounding effects of allotony. While McWhorter does not address this latter point directly, he mentions the general impact of allophony on the metric. He recognizes that a relatively small phonemic inventory may yield more allophonic contrasts than a relatively expansive one that is little affected by allophonic variation. McWhorter points out that it is the latter grammar, not the former, which will evaluate as complex by his standards.

[2] McWhorter unfortunately refers to the third module separately as "grammar" when his intent is to measure the complexity of "grammar" in a sense that includes all four components at once.

By way of providing concrete examples of simple and complex phonologies, McWhorter (2001: 138–139) compares the phonemic inventories of Tsez (Nakh-Daghestanian) and Saramaccan creole. Tsez includes 42 segments including uvular, pharyngeal, and pharyngealized uvular consonants as well as a phonemically long low central vowel. Furthermore, he reports that most stops and affricates have phonemic ejective counterparts. Tsez consonants generally have labialized phonemic variants, as well. By contrast, McWhorter claims that Saramaccan has a phonemic inventory of 25 segments (revised to 30 in McWhorter 2005: 48), only five of which are considered marked, namely, three prenasalized and two prevelarized consonants. Saramaccan has a limited tense/lax distinction in mid vowels and Tsez has a phonemic length distinction for the low central vowel /a/. It is not clear, based on McWhorter's description of the vowel system alone, which of these languages has the more marked system. Nonetheless, it is clear that he has chosen to emphasize the relative markedness of the consonant inventories in order to highlight the relative complexity of Tsez.

Arends (2001) critically observes that McWhorter's definitions of complexity are purely quantitative: more of some linguistic element equates directly with more complexity. Hence, more (marked) phonemes, more inflectional morphology, more syntactic rules, and/or more semantic distinctions all result in a grammar that is judged to be more complex. Arends contends that qualitative features, such as the internal complexity of rules, ought to be taken into account. Furthermore, he maintains that processes, such as phonological contraction rules, are given short shrift in McWhorter's methodology, while grammatical elements such as phonemes and tones are overemphasized. Arends also opines that it ought to be possible to weight rules according to their complexity. He poses the conundrum, "Which grammar is more complex, the one with *n* rules, each of complexity *C*, or the one with 2*n* rules, each of complexity C/2?" (181).

Although he vigorously challenged McWhorter's metric of complexity and proposed his own based on the biological notion of "dominance," Hagège (2001) noted that McWhorter's diagnostic for phonological complexity, predicated on the number of marked phonemes, is "the least controversial" (168). Kusters & Muysken (2001) observe that McWhorter's assessment of phonological complexity could be improved by taking into account distinctive features rather than raw phoneme counts: The relatively large phonemic set /p t k p^h t^h k^h b d g b^h d^h g^h/ can be reduced to four binary distinctive features while the smaller inventory of /p t k s d k^h/ requires five distinctive features to be analyzed parsimoniously. Hence, Kusters & Muysken wonder which system is more complex, the one that contains more phonemes, or the one that requires more distinctive features for analyzing their relationships.

Wurzel (2001) makes a number of insightful comments on McWhorter's phonological diagnostic. Since it is the only measure in McWhorter's methodology that involves the notion of markedness, Wurzel questions how McWhorter can possibly claim that there are some languages with only unmarked sounds, since all sounds are marked relative to one another in any given language, with /a/ as the best candidate for a fully unmarked element. No language, Wurzel points out, uses only the low central vowel, so even those languages with three-vowel systems have two vocalic counterparts juxtaposed to /a/ in degree of markedness. Thus, he calls on McWhorter to approach markedness as a scalar rather than a binary property of phonemes. It appears that McWhorter labels as simply "marked" those sounds that Wurzel would prefer to call "strongly marked" (ejectives, clicks, labialized consonants, etc.). He also criticizes McWhorter's phonological metric for its lack of any higher-order analysis other than that of phoneme inventories. Wurzel (2001: 379) quite reasonably remarks that "languages would seem simpler when they are limited to CV syllables than when they permit CCCVCCC syllables."

1.2.2. Complexity of inflectional morphology

In the realm of inflectional morphology, McWhorter argues that older, non-creole languages tend towards systems in which grammatical information is encoded in words via morphophonemic alternations, allomorphy, and suppletion. For example, he compares the different strategies for encoding the phrase 'I wanted' in Swahili *ni-li-taka* (1SG-PAST-want) and Mandarin *wǒ yào le* (1SG want PERF). In Swahili, the morphemes representing 1SG and PAST are phonetically incorporated with the head of the phrase, the verb *taka*. Conversely, in Mandarin the morphemes are quite free. McWhorter (2001: 137) rightly observes that "there is no a priori reason to assume that inflectional encoding is inherently more complex than encoding the same feature with a free morpheme." However, he goes on to assert that inflection usually winds up having "wider repercussions…which are complexifying factors in terms of exerting a load upon processing" (ibid.). While morphophonological alternations are often phonetically motivated (e.g. the addition of the allomorph [ɨz] instead of [s] or [z] to most English words that end in sibilant consonants), others (McWhorter cites consonant gradation in in Welsh: *ei gath* 'his cat' versus *ei chath* 'her cat') appear to fall outside the boundaries of natural phonological rules or constraints. Phonetically unnatural rules must be stored in the speaker-listener's memory, a task that McWhorter equates with greater complexity.

"Declensional and arbitrary allomorphy," according to McWhorter, also adds complexity to a grammar. He agrees that there is no reason to believe that a Latin synthetic dative construction such as *puero* is by itself any more complex than the comparable analytic English construction *to the boy*. McWhorter argues that Latin's system is more complex not because Latin nouns have case-sensitive desinences and English has prepositions, but because Latin nouns can belong to any of five inflectional classes, each with a different case paradigm. Thus, inflectional morphology triggers secondary complications in grammars. Furthermore, McWhorter states that greater complexity is to be found in languages where a given alternation can be expressed by using a number of morphological strategies that all serve the same synchronic purpose. For example, in Russian, aspectual distinctions are signaled by the prefixes *na-* and *vy-* (among others) depending on the stem—an alternation that apparently must be stored alongside the stem in the speaker-listener's mental lexicon.

In sum, McWhorter (2001: 138) claims that "inflection always complexifies grammar" based on (1) "the effects that inflection typically has upon a grammar over time and (2) the fact that some inflection, such as gender marking and declensional noun classes, does not correspond to concepts expressed by all grammars, but is instead purely supplementary to a grammar's machinery."

The criticism of McWhorter's inflectional morphology metric was not particularly severe and tended to focus on the theoretical basis for claiming that a given morphological strategy is more complex than another. Hagège (2001: 169) wonders, "[I]n exactly what respect can we say that [McWhorter's] examples of inflectional morphology…are more complex than the facts found in languages without inflectional morphology?"

On the other hand, Wurzel (2001: 380) actually makes a stronger claim than McWhorter's, arguing that "morphological symbolization" may in fact be more complex than the analytic constructions (e.g. English *to the boy*) because it results in more complex word forms. He cites the following example:

(1) Turkish (Turkic)
dol-dur-ma-yabil-ir-di-m
fill(itself)-CAUS-NEG-IMPOSS-AOR-PRET-1SG
'I could have refrained from filling (it/something) in.'

In disagreement with Hagège (2001), Wurzel opines that "extreme morphological complexity of word forms as such, as found with agglutination and especially incorporation, might contribute to the complexity of the language as a whole" (2001: 380).

To those who criticize McWhorter's decision to exclude derivational morphology from the complexity metric, he responds that derivational morphology has a higher degree of functionality than inflectional morphology, thus making it more operative than the optional linguistic "baggage" that he rates as complex. As an example, McWhorter (2001a) cites causativity. It would appear that all languages need to express the notion of causativity and this is often done using derivational morphology (though of course causativity may be expressed lexically as well). McWhorter further claims that all languages have some capacity for morphological derivation.[3] Thus, causative derivational morphology is judged to be more functional than inflectional morphology, which McWhorter characterizes as "decorative ornament" amassed by generations of speaker-listeners of the world's older languages.

2.0. Methods

2.1. Complexity metrics

2.1.1. Phonology

The present investigation will elaborate on McWhorter's (2001) phonological complexity metric and will thereby seek to avoid some of the criticism directed at the original version. McWhorter certainly

[3] In a personal communication, the author notes that this should only be taken to mean that all languages have derivational processes, not affixation in the strict sense.

proceeded in the least controversial direction by using phonemic inventories instead of phonological rules as the basis of his diagnostic. Given the spotty descriptions of so many of the world's phonologies, it is difficult to imagine a study in which the number of phonological rules or constraints could be listed exhaustively for each and agreed upon by even a quorum of phonologists, let alone weighted for their relative degrees of complexity (Arends 2001: 181).

Wurzel (2001: 379) makes a good point by protesting McWhorter's failure to include syllabic structure in his diagnostic. It seems rather shortsighted to leave such factors as consonant clusters and coda constraints outside the realm of phonological complexity. Maddieson (1984) used information about phonemic inventories, tonal distinctions, and attested syllable types as a way of calculating the number of possible syllables for any given language. Everett & Kern (1997) used this method in their grammar of Wari' (Chapacuara-Wanham, Madeira). The present study will make use of the same method. The result for each language will represent the total number of possible syllables, i.e. the combinatory possibilities of all syllable types (CV, CVC, etc.) based on the number of vowels, consonants, tones, diphthongs, laryngeal contrasts, and length contrasts reported in each phonological inventory.

I report the number of vowels as the number of monophthongs multiplied by length contrasts (e.g. in Hausa, Squamish, Wichita, and Nandi). Among the languages in the sample, voice quality is contrastive for vowels in Ju|'hoan, Hmong Njua, and Burmese. For Ju|'hoan, I calculated the number of vowels based on voice quality contrasts. For Burmese and Hmong Njua, however, where laryngeal qualities of vowels (e.g. breathiness) are associated exclusively with certain tonal contrasts (e.g. falling breathy tone) I use the number of basic vowels (those without laryngeal contrasts) and multiply them by an appropriate number of tones. Diphthongs are added to the number of vowels before calculating the number of potential syllables. This quantitative measure will be taken as a proxy for the more intuitive notion of phonological complexity.

One challenge to my algorithm as heretofore stated arises from the phonotactic restrictions on the syllable. Just because the shape CCVC occurs in English does not mean that the initial cluster may be comprised of all possible pairings of English consonants (as presumed by simply squaring the number of consonants for the onset cluster). Otherwise, Chomsky and Halle's (1968) famously monstrous form **bnick* would be starred no more. Likewise, many languages restrict the consonants that may occur in coda position. Thus, it would be incorrect to use the same number of consonants in onset position as in coda position when only a subset of the consonants can occur in the latter. Brazilian Portuguese, for example, allows only continuants in coda position, thus limiting the number of combinatory possibilities for syllable type CVC and for the range of possible syllables in general. In the languages I surveyed, the authors of the individual grammars described phonotactic restrictions on the syllable, if there were any. All phonotactic restrictions considered in the relevant sources are described under the heading of each language in §3.1.

A brief illustration of the algorithm follows: Imagine that Language X has three phonemic consonants, /p t k/ and three vowels /i u a/. Furthermore, imagine that in this language there are four syllable types: V, VC, CV, and CVC. Assuming no distributional restrictions, the number of possible syllables in Language X can be calculated by the following formula, where C and V represent the number of consonants and vowels: $V+(V*C)+(C*V)+(C*V*C)=3+9+9+27=48$.

Suppose we encounter another language, Y, with the same phonemic inventory, but one in which /p/ is banned in coda position. This reduces the total number of possible VC and CVC syllables in Language Y, since these are the only syllabic structures with coda consonants. For these syllable types, final C=2. Thus, the equation for total number of possible syllables yields 36.

Matters become more complicated with the addition of new syllable types and phonotactic constraints. Let us imagine that Language Z, in addition to the consonant inventories of X and Y, has a liquid phoneme /r/ which can occur as the first element in final clusters and as the second element in initial clusters. Further, clusters are only admissible if /r/ is one of the elements therein. Syllable types for Language Z include V, VC, CV, CVC, CCV, VCC, CCVC, CVCC, and CCVCC. Thus, if there were no distributional restrictions (e.g. if words like [rkit] were allowed), the result of the formula would look like this: $V+(V*C)+(C*V)+(C*V*C)+(C*C*V)+(V*C*C)+(C*C*V*C)+(C*V*C*C)+(C*C*V*C*C) = 3+12+12+48+48+48+192+192+768=1{,}323$.

However, this result admits words like /rkit/ and /ptat/, which the previously mentioned phonotactic rules disallow. For clusters, one of the consonants must be /r/, so the value for one of the C's must equal 1 (of course, for mathematical purposes, it doesn't matter which one), whereas the value for the other consonant is 4, i.e. it can vary freely among all four consonants (this assumes that geminate /r/ is possible in Language Z). Hence, the calculations for (C*C*V)=12; (V*C*C)=12; (C*C*V*C)=48; (C*V*C*C)=48; (C*C*V*C*C)=192. The total number of possible syllables for Language Z is thereby greatly reduced because of the cluster restrictions, from the initial calculation of 1,323 to 459.

The reason for calculating the number of possible syllables is to provide an objective, comparative criterion for measuring the combinatory potential of a language's phonological system. Admittedly, my algorithmic approach churns out many syllables that a language may not happen to incorporate in its lexicon. Thus, implicitly, I refer to phonological complexity *in potentia*, not complexity *in praesentia* (as suggested by an anonymous reviewer). I will now outline my reasons for doing so.

The number of potential syllables in a language is dictated by its phonemic contrasts, syllable types, and various phonotactic constraints. Given a discrete phonological inventory, there must be a discrete maximum syllable count. The actual syllable count, if it were of interest, could be derived from the maximum syllable count minus any number of accidental gaps (potential syllables that do not occur). This would constitute, in effect, a database of lexical syllables. Absent any means of automatically calculating accidental gaps, actual syllable counts can only be tabulated through exhaustive lexical research (i.e. counting). However, it is not only convenient but also reasonable to situate accidental gaps in the lexical domain and thus disregard them for present purposes. Accidental gaps could potentially be filled through innovative processes such as borrowing or coinage, which are not strictly dependent on phonology. For example, if the semantic content of a foreign word is useful to a community of speakers, it is not likely to be rejected by them because it contains an unheard-of syllable. (This is different than saying a word must be remodeled to fit native syllabic *structures*, which of course is quite common). Since sufficient phonological complexity already exists to accommodate new words, a measure of potential phonological complexity does not represent an alternate phonological reality. In fact, it comprises numerous (existing and non-existing) lexical realities. A measure of potential phonological complexity takes into account the fact that a relatively complex phonology (i.e. one with more potential syllables) can admit new words containing potential though unrealized syllables while a relatively less complex phonology must restructure words containing impossible, unrealized syllables.

While measuring potential phonological complexity is admittedly a different enterprise than measuring actual lexical complexity, it is nevertheless a reasonable choice to make when attempting to draw correlations using phonology (the combinatory possibilities which *can be* derived from the sound system) rather than the lexicon (the combinatory possibilities currently derived from the sound system and associated with meaning). Both lexical-morphological and phonological-morphological tests of the negative correlation hypothesis are valid; I have explicitly chosen to pursue the latter.[4]

As an anonymous reviewer points out, one must distinguish between complexity as it is experienced by the speaker and complexity as it is experienced by the listener, since there is no a priori reason the two should be equivalent. The reviewer observes that field-workers and those with a deep understanding of low-synthesis languages are aware of the fact that when a language is low on grammatical encoding, hearers must employ more "pragmatic machinery" to retrieve the speaker's intention. Thus, it is possible that a complexity equilibrium is reached between speaker and hearer. In the present study, my complexity measures are indeed locutocentric, for practical reasons: it is unclear how to determine how much cognitive effort a hearer might expend in processing a linguistic signal, whether in distinguishing between [pa] and [p'a] or between morphemes corresponding to passive and active voice. One way of sampling the complexity experienced by the hearer might be to measure the hearers' neural responses to individual linguistic stimuli and somehow correlate this measure with the effort required to disambiguate an utterance. It could be considered whether increased contextualization of utterances (or the ratcheting up of other pragmatic machinery) reduces this measure of effort. Experiments of this type could be of profound interest, but would certainly abound with complex empirical issues best resolved by cognitive scientists.

---

[4] Neither should it be presumed, as suggested by one anonymous reviewer, that my phonological complexity metric is all potential complexity while AUTOTYP's inflectional synthesis score (§2.1.2) is based entirely on actual complexity. Nichols (personal communication) indicates that her linguistic informants who speak languages with high inflectional synthesis scores often expressed disapproval at verbal forms including more than four morphemes, though in theory the verbs had valence to spare.

At this point, it will be useful to address the matter of diphthongs. According to my metric, even a small number of diphthongs will greatly multiply a language's phonological complexity. When computing the total number of potential syllables, the number of diphthongs is added to the number of vowels because of their shared role as syllabic nuclei. This method of dealing with diphthongs presents two difficulties. First, in languages with consonant-final syllables that incorporate glides in their consonant inventories, my algorithm will generate sequences such as [auw] and [aij] where the off-glide of the diphthong is followed by a full glide. None of the grammars in this study argued for tautosyllabic geminate glides. Thus, an appropriate number of syllables have been deducted on a case-by-case basis from the count of potential syllables for languages with all of the following three characteristics: (1) consonant-final syllable types; (2) glides in the consonant inventory; and (3) diphthongs. Yimas (Lower Sepik), Somali, and Korean are examples of such languages.

To illustrate the algorithm for computing phonological complexity, let us take a closer look at Lakhota, a Siouan language of North America. Lakhota has eight vowels and 17 consonants; it permits no diphthongs (Carter 1974: 155). There are five different syllable types, viz., V, CV, CVC, CCV, and CCVC. Carter exhaustively lists all 49 possible syllable-initial consonant clusters. So, to calculate the total number of potential syllables (σ_p), we need only multiply the number of vowels by consonants and sum the products for each syllable type: 8+(17*8)+(17*8*17)+(49*8)+(49*8*17)=9,512.

The algorithm may be formalized in the following equation, where *n* is the number of syllable structures and φ represents the product of the number of phonemic consonants, vowels (including length), diphthongs (D), tones (T), and laryngeal contrasts (LC) for the *i*th syllable type:

$$\sum_{i=1}^{n} \varphi_i \text{ , where } \varphi = (\ldots C)*(V+D)*T*LC*(C\ldots).$$

2.1.2. Morphology (inflectional)

The primary metric for measuring morphological complexity in this study will be inflectional synthesis as defined by Bickel & Nichols (2005) and as applied to a wide variety of languages in the ongoing AUTOTYP project (2004). Since the authors have generously made AUTOTYP results available for this study, it is necessary to describe the principles and assumptions underlying the collection and categorization of their data. This will involve both a brief explanation of the theoretical principles of morphology and a discussion of how these principles are implemented in AUTOTYP.

AUTOTYP's measure of inflectional synthesis deals with the degree to which verbs can be marked by inflectional categories. Because I use AUTOTYP data in the present study, I accept Bickel & Nichols' (2005) definition of inflection for practical reasons. For them, inflectional categories refer to grammatical categories "whose presence or shape is (at least in part) a regular response to the grammatical environment." They go on to list their "prime candidates" as "agreement, tense/aspect/mood, evidentials/miratives, status (realis, irrealis, etc.), polarity (negation), illocution (interrogative, declarative, imperative), and voice (including Austronesian-style verb orientation)."

Truly derivational operations—definitionally contrastive with inflectional operations—are known to change (1) the category of a root; (2) the valence (transitivity) of a verb root; or (3) the basic concept expressed by the root (e.g. distributives and diminutives). Unlike derivational categories, as mentioned earlier, inflectional categories are usually sensitive to the syntactic environment in which they occur. Bickel & Nichols (2005) illustrate: NPs have agreement relations with verbs, tense marking involves sequencing of tense rules, and cross-clausal anaphora allows distinctions in voice. Often, however, the grammatical sensitivity seems more narrowly morphological than syntactic. For example, different tense forms or negative versus affirmative forms may invoke different morphological paradigms, or combine with different sets of aspect forms or voices. While information about morphological sensitivity can usually be gleaned from basic descriptions, data on syntactic sensitivity is often unavailable. Bickel & Nichols (2005) therefore made the default assumption that, unless there is positive evidence for any grammatical sensitivity of the category, it is not considered inflectional but instead derivational. Spencer (1998: 9) cautiously observes, however, that distinguishing inflection and derivation "in such a way that it gives sensible answers for all languages" is an extremely difficult task.

Having described what is meant by "inflectional," I now turn to the definition of "synthesis," as formulated by Bickel & Nichols (2005). Certain grammatical categories (e.g. tense, voice, and agreement) can be expressed either by individual words or by affixes attached to some other word (or the stem of a word). "Synthetic" means that morphemes are combined for this purpose. 'Analytic' means that the morphemes are left uncombined. Bickel & Nichols (2005) provide the following example: In English, the auxiliary + verb sequence 'will wash' is analytic but the combination of verb and morphological past marker 'washed' is synthetic. How do we know, however, that 'will-wash' is not a single synthetic word? The following facts are revealing: (1) 'will' can appear as its own word, unlike '-ed', e.g. 'If you don't wash the dishes, I will'; (2) the position of 'will' relative to 'wash' can change, e.g. 'Will you wash the dishes?'; and (3) 'will' can be focused or emphasized, e.g. 'You *will* wash the dishes.' Bickel & Nichols (2005) make another default assumption "that phonologically separate particles are also syntactically separate, i.e. that they do not synthesize." However, when these particles cannot be used alone (i.e. they are morphologically dependent) or when they cannot move freely away from other words (in languages with relatively free word order) this is taken as evidence of synthesis.

Comrie's (1989) index of synthesis delineates the number of morphemes per word, thereby creating a continuum of languages with so-called "isolating" and "polysynthetic" languages at the two poles. Payne (1997: 27) believes that "the Chinese languages" come close to the isolating extreme and that Inuit falls at the polysynthetic end of the spectrum. He cites the following morphologically complex structure:

(2) Inuit (Eskimo, Eskimo-Aleut)
tuntussurqatarniksaitengqiggtuq
tuntu-ssur-qatar-ni-ksaite-ngqiggte-uq
reindeer-hunt-FUT-say-NEG-again-3SG:IND
'He had not yet said again that he was going to hunt reindeer.'

Polysynthetic languages are also traditionally graded according to the degree of fusion that occurs among the morphemes that combine to form words. Bybee (1985: 4) claims that the degree of morphological fusion of an affix to a stem "correlates with the degree of semantic relevance of the affix to the stem." Among the synthetic languages, some possess words that are easily dissected into constituent morphemes. In this category of languages, traditionally called "agglutinative," morphological and phonological boundaries (especially syllable boundaries) coincide. In non-agglutinative or fusional languages, there are various forms of assimilation, sandhi, allomorphy, and reduction at morpheme boundaries, which lead to more opacity in segmentation. Payne (1997: 28) refers to examples in English that demonstrate both a high and low degree of fusion. For example, when conveying the past tense in "strong" verbs like 'sang,' 'brought,' and 'thought' the quality of an internal vowel is modified. This contrasts with an English word-building process in which constituents can be straightforwardly segmented into meaningful units, e.g. 'anti-dis-establish-ment-ari-an-ism.'

Thus, inflectional synthesis avails us of hard evidence regarding the number of categories that can be morphologically synthesized on the verb. This is a narrowly-defined criterion for determining the morphological complexity of a language, but it is admittedly only a snapshot of a wider panorama. First of all, inflectional synthesis excludes, by definition, the many derivational or category-changing processes that can occur in a language. For example, according to AUTOTYP, Greenlandic Eskimo (Eskimo, Eskimo-Aleut) has a lower than mean inflectional synthesis score of 3 (on a normalized scale of 1-7), while it is well known for its rich derivational morphology. On the other hand, Koasati (Muskogean), with a high inflectional synthesis score of 7 (again on a normalized scale of 1–7) also demonstrates numerous derivational processes, e.g. a total of nine methods by which verbs may be derived from nouns (Kimball 1991: 335–353). AUTOTYP's inflectional synthesis measure tells us nothing about nominal paradigmatic complexity. In fact, the measure excludes, by definition, all non-verbal morphology. While the exclusion of morphological derivation limits my analysis and conclusions, the limitation is a necessary one until we have access to derivational data comparable (in terms of genetic diversity, systematicity, and quantifiability) to the verbal inflection data now residing in AUTOTYP. Furthermore, as McWhorter (personal communication) points out, it is reasonable to sample inflectional morphology specifically because of its "usually richer and more widespread interaction with syntax, this interaction being of note in covering the general issue of complexity more widely."

Bickel & Nichols (2005) explain how they measured inflectional synthesis for each language in the AUTOTYP sample. The elements under analysis were the "maximally inflected verb forms," i.e. the verbal form that is most synthetic. The number of inflectional categories that could be attached to the verb constituted the categories per word (cpw). For example, English, which has verbs that express person agreement (present =-s) and tense (past =-ed), has an inflectional synthesis degree of 2 cpw. Vietnamese, where there was no evidence of synthetic inflection on the verb, has 0 cpw. Koasati (Muskogean), at the opposite end of the spectrum, has inflected verb forms where the degree of synthesis can reach 13 cpw. When the same or similar category (e.g. number- or pluractional-marking agreement) appears at sufficiently distinct, multiple places in the verb, the category was counted twice. This is how Koasati's final inflectional synthesis score came to be 28 instead of 13 cpw.

Just as categories could be counted twice if they appeared in various positions, semantically-related categories were collapsed if they appeared in the same position. For example, tense, aspect, and mood were counted as one category unless they were clearly distributed in distinct positions. If categories such as agreement and tense (e.g. in the German suffixes–e and–en) cumulated into a single formative (cumulative exponence), the categories were counted separately. Furthermore, agreement was counted as one category per role (e.g. subject agreement, object agreement, etc.) regardless of the number of features that required agreement on the verb.

Due to these decisions, which often tend to disfavor synthesis, Bickel & Nichols (2005) admit that categories per word may be judged lower in languages that have been described less thoroughly. However, they state that their "survey showed that there is no correlation between obtained cpw value and the quality of description." Individual cpw values are in many cases valid within a confidence interval of ±1.

In the entire AUTOTYP sample, the most common scores are between 4 and 8, which accounts for 2/3 of the entire sample with a mode of 4 (17.4%). The distribution is uneven geographically, with a preponderance of high-synthesis languages in the Americas and low-synthesis languages in Eurasia.

While McWhorter (2001a) was reluctant to state that inflectional morphology is by itself more complex than analytic (he averred that inflectional morphology was simply indicative of more complex paradigms and memory-taxing allomorphs), I accept Wurzel's stronger view that complex inflectional morphology "might contribute to the complexity of a language as a whole" (2001: 380).

2.2. The survey

A cursory glance at one or two languages may appear to prove or disprove the negative correlation hypothesis, depending on the languages in question. For example, Koiari (Central and Southeastern, Trans-New Guinea) has the lowest degree of phonological complexity (70 potential syllables or a standardized Z-score of -0.65) in the sample and one of the lowest inflectional synthesis scores in AUTOTYP (Z=-0.89). Vietnamese, on the other hand, has a relatively high level of phonological complexity (Z=1.69) even though it earned the lowest inflectional synthesis score of all the languages in AUTOTYP (Z=-1.72). Vietnamese seems to support the negative correlation hypothesis while Koiari appears to challenge it. Thus, we must make every effort to sample a cohort of the world's languages that adequately represents their diversity.

Linguists have given a good deal of consideration to the manner in which individual languages should be chosen in order to produce a reliable snapshot of all languages for typological purposes (e.g. Bybee 1985; Nichols 1992; Bybee et al. 1994; Whaley 1997; Rijkhoff & Bakker 1998). Rijkhoff & Bakker (1998: 264–268) summarize the types of samples that are generally implemented in typological studies, the kinds of questions they propose to answer, and the weaknesses they harbor. Rijkhoff & Bakker distinguish between three types of samples: probability, random, and variety.

Probability samples are typically used to investigate correlations between the occurrence of linguistic phenomena or to calculate the probability of occurrence of a specific phenomenon. It is imperative that the languages in a probability sample are not of the same genetic origin. Counting languages from a common source in a probability sample is akin to counting the same language twice, since related languages share characteristics of their common ancestor. Other biases to avoid are cultural and areal—geographically or culturally connected languages, though genetically distinct, may share the linguistic properties that are under scrutiny. There is something of a paradox involved in constructing a probability sample, namely, if a sample is sufficiently large, it is "practically impossible" to exclude languages that are genetically related or spoken in the same region (Rijkhoff & Bakker 1998: 265). If a sample is too small, it will come up short in representing the diversity of the world's languages.

Conversely, a random sample is designed to ignore any form of genetic, typological, geographic, or cultural stratification. Of course, such a sample must be relatively large in order to generate reliable results.

Finally, a variety sample, like the one detailed by Rijkhoff & Bakker (1998) is designed to display the greatest possible linguistic variety. Rijkhoff & Bakker state that the purpose of variety samples is to conduct exploratory research, when little is known about the linguistic parameter under investigation and the sample must therefore include a wide variety of languages to ensure that its presence or absence is reliably captured. Variety samples take into account information about genetic relatedness and the degree to which language families are internally diverse (i.e. the number of branches below a certain node in the family's cladogram).

In a random sample, all the world's languages have an equal chance of being selected. By imposing restrictions on which languages can be sampled (e.g. excluding one of a pair that share a common ancestor), there will be more variation in the probability and variety samples. The results of a probability or variety sample might be considered more cautious than those of a random sample from the perspective of a linguist, since there is less likelihood of sampling multiple related languages that all manifest the same typological feature (due to genetic relation). From a statistician's point of view, however, the most conservative sampling technique is in fact random,[5] even though variety may decrease the chances of a spurious correlation. The same mathematical techniques need not be applied to a random sample on the one hand and a probability or variety sample on the other "as long as it is understood that the results obtained are a conservative test of the hypothesis" (Bybee et al. 1994: 304).

Here I have chosen a variety sample, as this is an exploratory study in which nothing is known a priori about the possible correlation between morphological and phonological complexity. Using the geographic and genetic classification schemes employed in AUTOTYP (Bickel & Nichols 2004), my guiding principles of selection were geographic distribution and descriptive adequacy (i.e. a substantial and reliable phonological description). At least one representative from each of AUTOTYP's narrow geographic areas was selected. In some cases, where the geographic area seemed to contain a great deal of linguistic diversity (i.e. a large number of families), the decision was made to sample from a number of those families. Thus, in the case of the North American Basin-Plains area, languages from three families, Caddoan, Kiowa-Tanoan, and Siouan were sampled. Similarly, two language families from Eastern North America were sampled (Muskogean and Na-Dene); two from Europe (Turkic and Indo-European); two from Inner New Guinea (Kalam and Lower Sepik); three from the "South African" region (Nilotic, Khoisan, and Benue-Congo); and three from Southeast Asia (Austroasiatic; Hmong-Mien; and Sino-Tibetan). As mentioned previously, the sample was also constrained by the necessity for an adequate phonological description of each language. For present purposes, an "adequate" phonological description consisted of (1) the number of phonemic consonants; (2) the number of phonemic vowels (including length distinctions and laryngeal contrasts); (3) the number of phonemic tones (if any); (4) the number and kind of syllable types; and (5) the phonotactic constraints on the syllable.

The following 32 languages (Table 1) are included in the study (their geographic distribution is depicted in Figures 1 and 2):

TABLE 1. Languages in the sample (N=32).

Language	Family	Area	Phonology source(s)
Amele	Madang	North Coast New Guinea	Roberts (1987)
Arabic (Cairene)	Semitic	Northern Africa	Watson (2002); Harrell (1957)
Burmese	Sino-Tibetan	Southeast Asia	Okell (1969)
Diyari	Pama-Nyungan	Southern Australia	Austin (1981)
Greek (modern)	Indo-European	Europe	Holton et al. (1997)

[5] Analyses involving random samples should be complemented at some point by a stratification of the sample for any known factors that could potentially skew or confound the results. Nonetheless, it is presumed that in a random sample, chance is the only factor governing the initial selection of items from a population.

Hausa	Chadic	Northern Savannah	Newman (2000)
Hmong Njua	Hmong-Mien	Southeast Asia	Mortensen (2004)
Jacaltec	Mayan	Central America	Craig (1977)
Ju|'hoan	Northern Khoisan	Southern Africa	Dickens (1994) & Miller-Ockhuizen (2003)
Kannada	Dravidian	South/Southeast Asia	Sridhar (1990)
Kiowa	Kiowa-Tanoan	North American Basin	Watkins (1984)
Koasati	Muskogean	Eastern North America	Kimball (1991)
Kobon	Kalam	Inner New Guinea	Davies (1981)
Koiari	Koiari	Southern New Guinea	Dutton (2003)
Korean	isolate	North Coast Asia	Sohn (1999)
Lakhota	Siouan	North American Basin	Carter (1974)
Mapudungun	isolate	Andean	Smeets (1989)
Mongolian (Khalka)	Mongolian	Inner Asia	Kullmann (1996)
Nandi	Nilotic	Southern Africa	Creider & Creider (2001)
Ngarinjin	Worrorran	Northern Australia	Coate & Oates (1970)
Slave	Na-Dene	Eastern North America	Rice (1989)
Somali	Cushitic	Ethiopian Plateau	Saeed (1999)
Squamish	Salishan	Alaska-Oregon	Kuipers (1967)
Trumai	isolate	Southeast South America	Guirardello (1999)
Tukang Besi	Austronesian	Oceania	Donohue (1999)
Turkish	Turkic	Europe	Comrie (1997)
Vietnamese	Austroasiatic	Southeast Asia	Thompson (1987)
Wari'	Chapakuran	Northeast South America	Everett & Kern (1997)
Wichita	Caddoan	North American Basin	Rood (1976; 1996)
Yimas	Lower Sepik	Inner New Guinea	Foley (1991)
Yoruba	Benue-Congo	Southern Africa	Awobuluyi (1978)
Yurok	Algic	California	Blevins (2003; in preparation)

3.0. Typological data

3.1. Phonology

3.1.1. Amele

Amele has 15 consonants and five vowels (Roberts 1987). There are five diphthongs. Syllable types include V, CV, VC, and CVC. There are no phonotactic restrictions on the syllable (though some consonants in certain positions are admittedly rare, e.g. /g͡b/ in coda and /w/ in onset preceding /u/).

3.1.2. Arabic (Cairene)

Cairene Arabic has 25 consonants, not including eight "marginal phonemes attested in the dialect or among certain speakers of the dialect" (Watson 2002: 20–21). The vowel inventory consists of three short vowels and five long ones. Only three syllable types occur in Cairene Arabic, namely CV, CVC, and CVCC. While Watson does not list phonotactic restrictions on the syllable, Harrell (1957: 32–43) gives an exhaustive account of final clusters in colloquial Egyptian Arabic, totaling 223.

3.1.3. Burmese

Burmese has 34 consonants (excluding /r/ from English and Pali loanwords), and 12 monophthongs (Okell 1969). There are four diphthongs and four lexical tones, only one of which occurs on syllables closed by /ʔ/. Glottal stop and four voiced nasal consonants occur in syllable final position. There are eight syllable types, viz., V, CV, CVC, CCV, CCVC, CCCV, CCVC, and CCCVC. C_2 in CC clusters must be either /j/ or /w/ (/j/ after nine consonants and /w/ after all except five). In CCC clusters, C_2 and C_3 must be /j/ and /w/, respectively.

3.1.4. Diyari

The consonant inventory of Diyari consists of 22 phonemes, including a full set of retroflex stops (Austin 1981). Diyari has three vowel phonemes. No clusters are permitted in the language and there are only two syllable types, CV and CVC. Only nasals (six), laterals (four), and the tap may occur in syllable-final position. Though there are phonotactic restrictions (no laterals or tap) on word-initial consonants, there appear to be no such constraints on syllabic onsets.

3.1.5. Greek (modern)

Greek has 15 consonantal and five vocalic phonemes. Syllable structures are V, VC, VCC, CVC, CVCC, CCVC, CCVCC, CCCVC, CCCVCC, CV, CCV, and CCCV. Only /n s j w/ can occur in coda position and only /ks/ and /fs/ are admissible final clusters. There are, however, a wealth of onset clusters, including 56 biconsonantal clusters and seven triconsonantal clusters (Holton et al. 1997: 12–13).

3.1.6. Hausa: Chadic

Standard Hausa has 32 consonant phonemes, ten vowels, and two diphthongs. Only two syllable types occur in the language: CV and CVC. Hausa has three contrastive lexical tones: two level tones, which occur on short monophthongs (five) and a falling tone which occurs only on long vowels and diphthongs (seven). The phoneme Newman (2000) characterizes as /fy/ does not occur in coda position.

3.1.7. Hmong Njua (Mong Leng)

Hmong Njua has 40 consonants including uvular aspirated, unaspirated, nasalized, and pre-nasalized stops, a retroflex affricate, an alveopalatal fricative, and a voiceless lateral (but excluding aspirated /s^h/ and an intermittently-appearing velar nasal in coda position) (Mortensen 2004). There are six oral and three nasal monophthongs. There are four diphthongs. Hmong Njua has two syllable types only: CV and CCV. Additionally, there are seven lexically contrastive contour tones. As for phonotactic restrictions, the high front vowel /i/ (both nasal and oral varieties) fails to occur after uvular stops. Furthermore, the only clusters are closed by laterals and initiated by the set of nine labial and velar consonants. The falling breathy tone never occurs in syllables with aspirated onsets and the high falling and falling creaky tones only occur in syllables opened by aspirated consonants.

3.1.8. Jacaltec

Jacaltec has 27 consonants and five vowels. Syllable types include V, VC, CVC, and CCVC. From my own review of Craig's (1977) data, it appears that in consonant clusters, C_1 is always from the set /s, ʃ, ʂ/ and C_2 is drawn from the set /p t k q ʔ m l w/. Thus, the combinatory possibilities for this initial cluster are limited to 24.

3.1.9. Ju|'hoan

Miller-Ockhuizen (2003) accounts for 93 consonants and 34 vowels (she includes epiglottalization, length, and nasality as distinctive features). There are four lexical tone distinctions and five diphthongs (Miller-Ockhuizen 2003: 75). Syllable types include V, CV, and CVC. The author imposes a fairly strong restriction on Ju|'hoan roots, which for present purposes will be treated as syllables. She claims that no two "guttural" features can appear in the same syllable for reasons of perceptual salience. This effectively excludes combinations such as [epiglottal] + [breathy], etc. With this constraint in place, the total number of possible syllables can only be calculated by summing the products of all allowable CV and CVC combinations. I surmise from Miller-Ockhuizen (2003: 108) and from the entries in Dickens' (1994) dictionary that only /ŋ/ and /m/ occur in coda position.

3.1.10. Kannada

The phoneme inventory of Kannada consists of 22 consonants and 10 vowels (excluding loans, like the aspirate series from Sanskrit). According to Sridhar (1990: 297–298), with the influx of loan words, the maximal syllable is now CCCVCCC and the minimal syllable is V. Sridhar notes, however, that consonant clusters are found only in the modern variety of Kannada, due to the presence of loan words. Since he asserts that consonant clusters are broken up by the insertion of vowels the present analysis will focus on the older syllable structures of Kannada which seem to mirror popular speech. The syllable structures (uninfluenced by the consideration of loans) are V, CV, VC, and CVC. Geminates in the language are never tautosyllabic.

3.1.11. Kiowa

Kiowa has 21 consonants, 24 vowels, eight diphthongs, and three lexical tones. According to Watkins (1984), syllable types are V, CV, and CVC. Initial consonant clusters exist, but are simply allophones of the sequence velar stop + low back vowel, i.e. /Ka/ → [kja], so there is no need to inflate the number of potential syllables by including CCV types. Further phonotactic constraints cited by Watkins (1984) are a ban on dental and alveolar consonants before /i/; a ban on velars and /j/ preceding /e/; and a ban on all vowels except /u/ following velars. Coda consonants are limited to a set of six. HL tone occurs only on long vowels (including diphthongs) or on sequences of vowel + resonant.

3.1.12. Koasati

Koasati has 15 phonemic consonants, six vowels (three basic vowels plus a duration contrast), and three tones. Syllables in Koasati include CV and CVC. Consonant clusters are permitted only across syllable boundaries. According to Kimball (1991: 35) "All vowel clusters are phonetically separated by a glottal stop." No phonotactic restrictions on the syllable are indicated.

3.1.13. Kobon

Kobon has 23 consonants (including four prenasalized stops that I will analyze as phonemic) and seven vowels (Davies 1981). There are also two diphthongs. Syllable structures include V, VC, CV, CVC, CCV. The three consonants /h w j/ are banned in coda position. There are no restrictions on onsets. There are only four (initial) consonant clusters.

3.1.14. Koiari

Koiari has five vowels and 13 consonants (Dutton 1996). The two syllable types are V and CV. There are no positional restrictions on the consonants. There are 22 possible "vowel clusters," but it is not clear whether Dutton (2003) considers them nucleic diphthongs. In an unrelated discussion, he categorizes the word /arai/ (no gloss provided) as trisyllabic, suggesting that /ai/ is in fact disyllabic. For lack of a clear statement as to their existence, diphthongs are not included in this analysis.

3.1.15. Korean: isolate

Korean has 21 consonants, ten vowels, and 11 diphthongs (Sohn 1999). Syllable structures include V, CV, CCV, VC, CVC, and CCVC. The second member of a consonant cluster must be one of the glides /w j/.

3.1.16. Lakhota

The consonant inventory of Lakhota consists of 17 members (Carter 1974: 110). Carter (1974: 50–51) argues that the aspirate and ejective series are bisegmental clusters, necessitating syllables of the type CCV-. The vocalic inventory, containing both nasal and oral segments, includes eight phonemes. Syllable types consist of V, CV, CVC, CCV and CCVC. Carter (1974: 52) exhaustively lists and exemplifies the 49 possible initial clusters (including the aspirates which he analyzes as C+/x/ and the ejectives which he regards as C+/ʔ/).

3.1.17. Mapudungun

Mapudungun has 19 consonants and six vowels. Smeets (1989) describes the syllabic structure and allowable clusters of roots and suffixes separately. I will combine the restrictions manifest in each subset, as the present study has no reason to tease apart phonological distinctions made in the sub-categories of a language's morphology. Allowable syllable types are: V, CV, VC, CVC, and CCV (the latter is found in suffixes only). Five stops and affricates are banned from initial position, bringing the total to 14. The phoneme /q/ does not occur after vowels and the vowel /ü/ does not occur initially. Only five consonant clusters occur in the onset.

3.1.18. Mongolian (Khalka)

According to Kullmann (1996), Mongolian has 21 consonants, 14 vowels, and five diphthongs. Palatalized consonants have not been analyzed as phonemic due to a lack of clear discussion of the subject. Mongolian has four syllabic structures: V, VC, CV, and CVC. Only nine consonants may occur in the syllable coda.

3.1.19. Nandi

The Nandi phonemic system consists of 13 consonants, 20 vowels (ATR and length distinctions are operative), and five tones, not all in surface contrast (Creider & Creider 2001). Specifically, L, HL, and LL tones occur on both long and short vowels, while H occurs only on short vowels and LH occurs only on long vowels. Syllable types are V, CV, VC, CVC, and CVCC. The only possible final cluster is /nk/.

3.1.20. Ngarinjin

Ngarinijin has 20 consonants (including three prenasalized stops in initial position) and five vowels. Syllable structures include V, CV, VC, and CVC. Coate and Oates (1970) mention no phonotactic constraints on the syllable.

3.1.21. Slave: Na-Dene

The variety considered is the Bearlake dialect. Slave has 34 consonants. There are a total of 16 vowels: six short oral vowels, five short nasal vowels, three long oral vowels, and two long nasal vowels. Slave also has two lexical tones. According to Rice (1989: 143), there are syllables of the types V, VC, CV, and CVC. Only /h j ʔ/ occur in coda and there are no restrictions on onset.

3.1.22. Somali

Somali has 22 phonemic consonants (Saeed 1999). There are 20 vowels, based on an ATR system and a duration contrast. There are, in addition, three lexical tones. Syllables types include V, CV, and CVC. The consonants /t k j/ do not occur in coda position. HL tone must occur on a long vowel or diphthong (of which there are five).

3.1.23. Squamish

According to Kuipers (1967: 21), the Squamish sound system "comprises 29 consonants, four vowels, and a feature of glottalization (glottal stop)." The glottal stop occurs only in contact with a vowel or a sonorant, either as ʔV, ʔR, Vʔ, or Rʔ. For present purposes, I will treat glottal stop as a consonant and thereby raise the total number of consonants to 30. A three-way length contrast is functional among vowels. There are additionally three diphthongs. Syllable types include V, CV, VC, CVC, VCC, and CVCC. Only final clusters of a sonorant and glottal stop (six total) are permitted.

3.1.24. Trumai

The phonemic inventory of Trumai consists of 23 consonants and six vowels. There are four syllable types: CV, CVC, V, and VC. Guirardello (1999) reports no phonotactic constraints on the syllable, but indicates that restrictions exist on the word.

3.1.25. Tukang Besi

Tukang Besi has 22 consonant and five vowel phonemes (Donohue 1999: 15–16), with syllables of the type V, CV. Though there are 25 possible "vowel clusters," it can be inferred from his discussion that these are not to be considered tautosyllabic (i.e. diphthongs).

3.1.26. Turkish

Comrie (1997) describes Turkish as having 23 consonants and eight vowels. Syllables types include V, CV, VC, VCC, CVC, and CVCC. Constraints on syllable final clusters dictate that C_1 must be a nasal or liquid and C_2 must be a stop. While observing that there are some Perso-Arabic loans with fricatives in syllable-final clusters, Comrie (1997: 891) avers that "most final consonant clusters in loanwords are not imported into Turkish in syllable-final position," meaning that epenthetic processes intervene. Voiced stops do not occur in final position.

3.1.27. Vietnamese

The phonemic inventory of Vietnamese consists of 19 consonants, and 14 vowels, including "semi-vowels" (Thompson 1987). In addition, there are six lexical tones. I have accounted for glottal stop (which is not counted among the consonants because of its strict association with tone) by using all six lexical tones, some of which are conditioned by the presence of glottal stop. Syllable types include CVC, CV, and V. There are 40 diphthongs (vowel + vowel; vowel + semi-vowel; and semi-vowel + vowel). Only eight consonants occur finally; 18 occur initially. The so-called "upper-vocalics" (mid to high vowels and semi-vowels) do not occur finally, unless they are the second element in a diphthong. /iê/ occurs before /w p m t n/; /uô/ occurs before /j m t n/; /u'o'/ occurs before all occurring codas; and /u'u̯'/ may not be followed by any consonant.

3.1.28. Wari'

Wari' has 15 consonant phonemes, six vowels, and eight diphthongs (all nasal) (Everett & Kern 1997). The language makes use of a unique voiceless apico-dental, bilabial vibrant (an allophone of /t/, so it is not included among the 15 phonemic consonants). Syllable types include CV and CVC. Five consonants may appear in coda and 13 in onset.

3.1.29. Wichita

Wichita has ten consonants and a total of nine vowels (a three-way length contrast is operative) (Rood 1976). Furthermore, there are two lexical tones. Based on Rood's (1996) assessment of the syllable canon as (C)(C)(C)(C)V(C)(C)(C)(C), there are 25 distinct syllable types. There are ten possible onset clusters (seven CC; one CCC; and one CCCC). There are numerous restrictions imposed on coda clusters. Rood (personal communication) cites only one CCCC cluster, /ncks/. Final CCC clusters may be composed of stop-continuant-stop, continuant-stop-continuant, or a series of continuants (Rood 1996). To these, Rood (personal communication) adds continuant-continuant-stop. Further phonotactic restrictions apply: CCC cannot include /r/; only /h s/ occur in sequences of continuant-stop-continuant; /hh/C is disallowed; geminates are restricted to /s:/ and /n:/; and there are no examples of "triplicate" clusters such as /hhh/ or /sss/. Final CC clusters may be alternating stop-continuant, continuant-stop, or continuant-continuant. The phoneme /k^w/ cannot occur in coda position. Furthermore, /y/ does not participate in surface clusters (Rood 1996).

3.1.30. Yimas

Yimas has 12 consonants, four vowels, and two diphthongs (Foley 1991). There are five syllable structures: V, CV, VC, CCV, and CVCC. Phonotactic constraints on the syllable allow only for initial clusters in which (1) C_1 is a non-palatal stop and C_2 is /r/; (2) C_1 is /p/ or /k/ and C_2 is /w/; or (3) C_1 is a nasal and C_2 is a homorganic stop. The only allowed coda cluster is one in which C_1 is a stop and C_2 is a homorganic nasal.

3.1.31. Yoruba

Yoruba has 23 consonants (including two labiovelar stops), 11 vowels, five diphthongs, and three lexical tones (Awobuluyi 1978). Syllable types include V and CV. Nasal consonants (three) may carry tone and function as syllabic nuclei.

3.1.32. Yurok

The consonant inventory of Yurok consists of 29 consonants and 11 vowels (a duration contrast is operative) (Blevins 2003: 135). There are 13 syllable types. All Yurok consonants occur alone in coda or onset position (though /x/ is attested finally in only three words). Blevins (in preparation) exhaustively lists 49 initial clusters (including both bi- and tri-consonantal types) and 27 final clusters. A total of 66 syllables are banned according to co-occurrence constraints, e.g. */wu yi uw iy ɚr/.

3.1.33. Summary

Table 2 summarizes the number of consonants (C), vowels (V), diphthongs (D), syllable types (σ), and potential syllables (σ_p) for the languages included in the survey. It also presents the log-transform (base 10) (Log-σ_p) of σ_p and Z-scores corresponding to σ_p and Log-σ_p. A discussion of the log transform is provided in §4.1.

TABLE 2. Consonants (C), vowels (V), diphthongs (D), syllable types (σ), and potential syllables (σ_p) for the languages in the sample. Log-transform (base 10) of σ_p (Log-σ_p) is given, as well as the standardized Z-scores of Log-σ_p.

Language	C	V	D	σ	σ_p	Log-σ_p	Z-Log-σ_p
Amele	15	5	6	4	2816	3.4496	-0.35626
Arabic (Cairene)	25	8	0	3	49800	4.6972	1.3989
Burmese	34	12	4	8	8384	3.9235	0.31033
Diyari	22	3	0	2	792	2.8987	-1.1313
Greek	15	5	6	9	5201	3.7161	0.018599
Hausa	32	10	2	2	4512	3.6544	-0.06823
Hmong Njua	46	9	4	2	4130	3.616	-0.12228
Jacaltec	27	5	0	4	3245	3.5112	-0.26962
Ju|'hoan	93	34	5	3	80316	4.9048	1.6909
Kannada	22	10	0	4	5290	3.7235	0.028966
Kiowa	21	24	8	3	10435	4.0185	0.44403
Koasati	15	6	0	2	4590	3.6618	-0.05776
Kobon	23	7	2	5	3653	3.5626	-0.19726
Koiari	13	5	0	2	70	1.8451	-2.6136
Korean	21	10	11	6	25134	4.4003	0.98112
Lakhota	17	8	0	5	9512	3.9783	0.38745

Mapudungun	19	6	0	4	1721	3.2358	-0.65711
Mongolian	21	14	5	4	4165	3.6196	-0.11712
Nandi	13	20	0	5	17730	4.2487	0.76791
Ngarinjin	20	5	0	4	1890	3.2765	-0.59988
Slave	34	16	0	4	4480	3.6513	-0.07258
Somali	22	20	5	3	4080	3.6107	-0.12972
Squamish	30	12	3	6	14865	4.1722	0.66022
Trumai	23	6	0	4	3456	3.5386	-0.23113
Tukang Besi	22	5	0	2	115	2.0607	-2.3103
Turkish	23	8	0	6	8256	3.9168	0.30093
Vietnamese	19	14	40	3	52007	4.7161	1.4254
Wari’	15	6	8	2	1926	3.2847	-0.58835
Wichita	10	9	0	25	79200	4.8987	1.6824
Yimas	12	4	2	5	884	2.9465	-1.0641
Yoruba	23	11	5	2	1161	3.0648	-0.89761
Yurok	29	11	0	13	48840	4.6888	1.387

Figure 1 illustrates the geographic distribution of phonological complexity across the globe. Diamonds indicate a relatively low number of potential syllables (one standard deviation or greater below the mean); circles indicate a medium number (between one standard deviation below and one standard deviation above the mean); and squares indicate a relatively high number (one standard deviation or greater above the mean).

FIGURE 1. Phonological complexity (Log-σ_p) in geographic distribution. Diamonds indicate languages more than one standard deviation below the mean (n=4). Circles indicate languages between -1 and 1 standard deviation from the mean (n=23). Squares indicate values more than one standard deviation above the mean (n=5).

Mapping the log-transform of σ_p suggests that in relative terms, phonologically simple languages, i.e. those one standard deviation or more below the mean, are found in Oceania and Australia. Again in relative terms, phonologically complex languages, i.e. those one or more standard deviations greater than the mean, are found in wider geographic dispersion, in Africa, Asia, and North America.

3.2. Morphology

Table 3 presents the statistical results of the AUTOTYP survey with regard to the relevant languages in the present sample. The inflectional synthesis (IS) value is a raw measure of categories per word (cpw), as described in §2.1.2.

TABLE 3. Inflectional synthesis (IS) scores along with standardized Z-scores of the same measure (Z-IS).

Language	IS	Z-IS
Wichita	29	2.29
Koasati	28	2.15
Ngarinjin	26	1.88
Lakhota	24	1.60
Mapudungun	19	0.91
Slave	19	0.91
Trumai	18	0.77
Yimas	16	0.49
Hausa	15	0.35
Squamish	15	0.35
Turkish	14	0.22
Yurok	14	0.22
Arabic (Cairene)	14	0.22
Korean	13	0.08
Tukang Besi	12	-0.06
Kiowa	12	-0.06
Amele	11	-0.20
Nandi	11	-0.20
Yoruba	10	-0.34
Jacaltec	10	-0.34
Kobon	10	-0.34
Wari'	7	-0.75
Hmong Njua	7	-0.75
Koiari	6	-0.89
Diyari	6	-0.89
Somali	6	-0.89
Mongolian	6	-0.89
Greek	6	-0.89
Kannada	6	-0.89
Burmese	6	-0.89
Ju|'hoan	2	-1.44
Vietnamese	0	-1.72

Figure 2 illustrates the geographic distribution of morphological complexity around the world. The shapes of the data points emphasize those languages that fall one standard deviation above and below the mean. There is apparently a tendency towards high morphological complexity in North America.

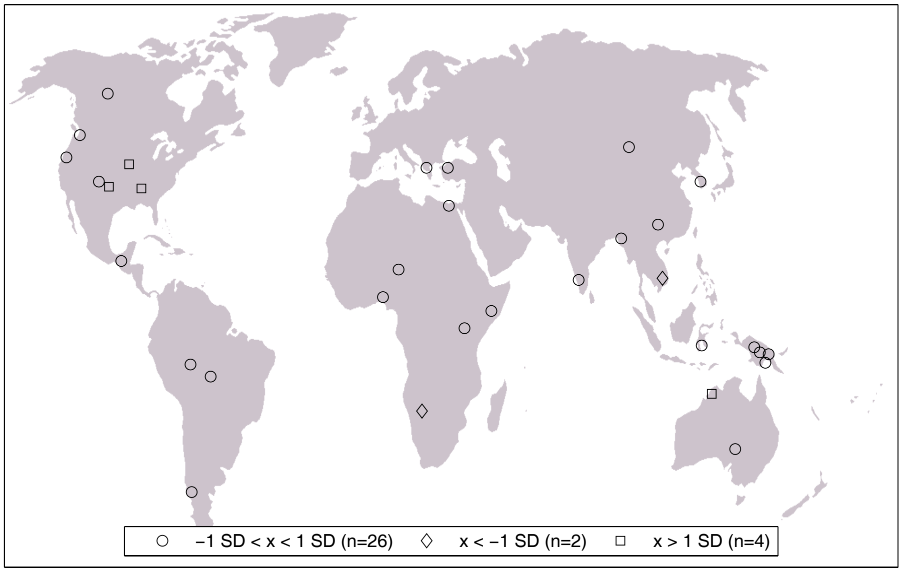

FIGURE 2. Inflectional synthesis (IS) on the verb in geographic distribution. Diamonds indicate languages more than one standard deviation below the mean (n=2). Circles indicate languages between -1 and 1 standard deviation from the mean (n=26). Squares indicate values more than one standard deviation above the mean (n=4).

4.0. Results

4.1. Descriptive results

The total number of languages in the sample is 32, representing 30 known families and two isolates across a wide geographic distribution (23 sub-areas in AUTOTYP's geographic classification system).

The distribution of languages across inflectional synthesis (IS) scores is skewed slightly to the right, with a fairly high standard deviation (M=12.44, SD=7.23) (see Figure 3).

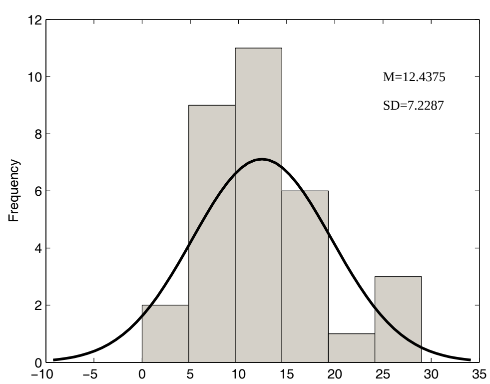

FIGURE 3. Histogram of inflectional synthesis scores (in cpw) for the 32-language sample.

However, the skewing is much more dramatic when it comes to the phonological complexity score (total number of potential syllables σ_p) for the languages in the sample. The histogram in Figure 4 clearly demonstrates that a few high scores exert a substantial outlier effect (M=14458, SD=22240.29).

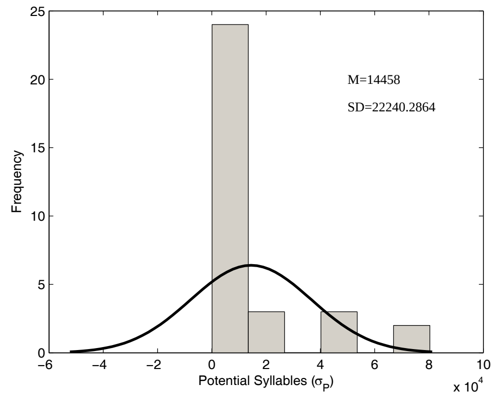

FIGURE 4. Histogram of potential syllable (σ_p) scores for the 32-language sample.The results of the Shapiro-Wilk test for normality agree with an intuitive assessment of Figure 4: the phonological scores are unlikely to come from a normal distribution ($W=0.635$, $p<0.0001$). By comparison, the IS scores are likely to come from a normal distribution ($W=0.937$, $p>0.05$).

The non-normality of the phonological scores makes it impossible to draw a correlation between independent variables which are assumed to be sampled from normal distributions. Following standard practice in dealing with right-skewed distributions (Hartwig & Dearing 1979), a logarithmic transformation (base 10) was applied to the σ_p values to generate something closer to a normal distribution, as presented in Figure 5. The mean of the log-transformed scores is approximately 3.7 with a standard deviation of 0.71, a great improvement over the skewed original scores. When applied to the log-transformed phonological scores, the Shapiro-Wilk test for normality indicates that the scores are most likely sampled from a normal distribution ($W=0.945$, $p>0.05$).

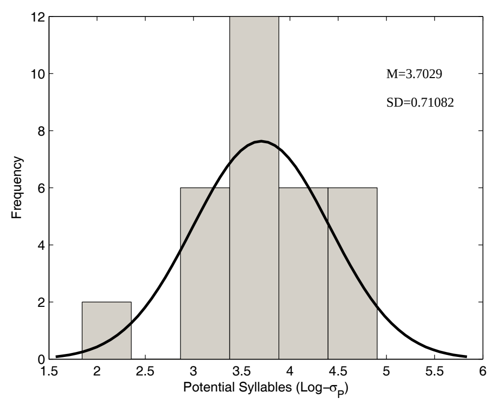

FIGURE 5. Histogram of the log-transform (base 10) of potential syllables (Log-σ_p) for the 32-language sample.

The logarithmic transformation reduces the relative spacing between the phonology scores on the right side of the distribution, where the average difference between the top four scores is 10,172 potential syllables. Since the transformation is logarithmic, the spacing of the left-hand scores in the distribution is less affected. On this side of the distribution, the distance between untransformed scores is much smaller (an average of 271 potential syllables between the four lowest). Thus, the logarithmic transform reduces non-normality while preserving the relative ordering of the scores. I rely on the log-transform value of σ_p (Log-σ_p) throughout the rest of the analysis because it is safe to presume that it comes from a normal distribution and can therefore be used in linear statistical models.

Histograms for the remaining measures will not be presented. However, Table 4 exhaustively reports the means and standard deviations for all the relevant variables: inflectional synthesis or categories per word (IS), number of consonants (C), vowels (V), diphthongs, syllable types (σ), potential syllables (σ_p), and the log-transform of potential syllables (Log-σ_p).

TABLE 4. Means (M) and standard deviations (SD) for inflectional synthesis (IS), phonemic consonants (C), phonemic vowels (V), diphthongs (D), syllable types (σ), and the log-transform of potential syllables (Log-σ_p) in a 32-language sample.

	IS	C	V	D	σ	Log-σ_p
M	12.438	24.25	10.25	3.625	4.875	3.7029
SD	7.2287	14.642	6.6478	7.3033	4.3681	0.71082

4.2. Correlation

Figure 6 illustrates the relationship between Log-σ_p (the primary metric of phonological complexity) and IS (the metric of morphological complexity) for the language sample. There is truly no discernible tendency, as statistical analysis shows: the Pearson product moment correlation (r) between IS and Log-σ_p amounts to only 0.0704, which fails to achieve significance at even the 0.05 level (p=0.702).

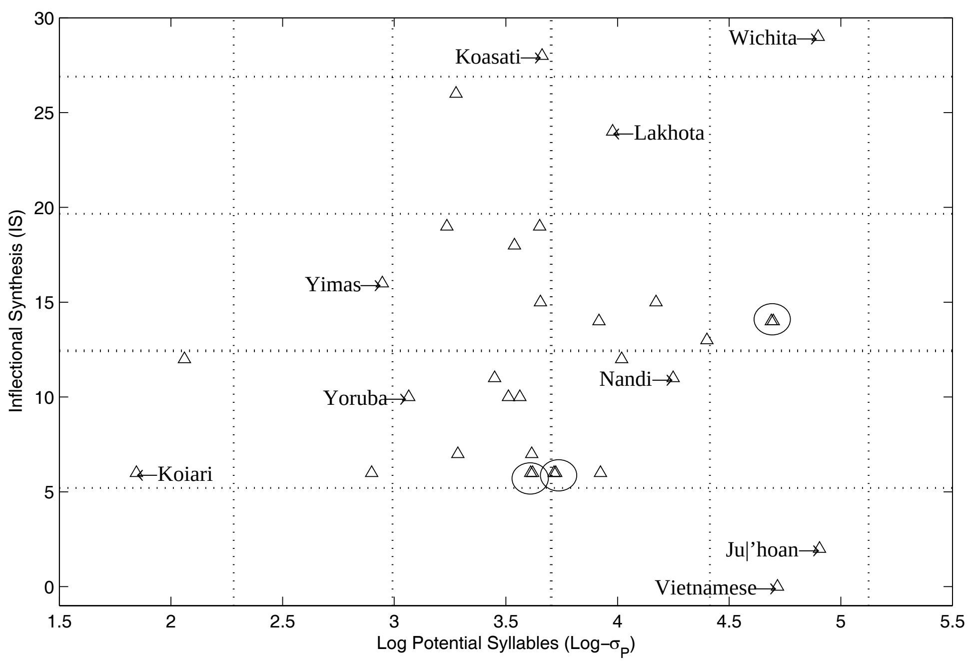

FIGURE 6. Phonological complexity (cardinal axis) and morphological complexity (ordinal axis), where trials equal languages (N=32). Phonological complexity is measured by the log-transform of potential syllables (Log-σ_p) and morphological complexity by inflectional synthesis or categories per word (IS). In three instances, two langugages are nearly identical in their scores; these clusters of two are each highlighted with circles. The bold vertical line represents mean Log-σ_p and the vertical dotted lines represent standard deviations from the mean. Likewise, the bold horizontal line represents mean IS and the horizontal dotted lines represent standard deviations from the mean. The Pearson product moment correlation (r) between IS and Log-σ_p is slightly positive (0.0704) but statistically insignificant ($p>0.05$).

For the sake of space and ease in reading, in the following discussion I will use the terms "simple" and "complex" as shorthand for the terms "relatively simple" and "relatively complex".[6] Figure 6 indicates that 11 languages (~34.4 %) are below the mean for IS and Log-σ_p, i.e. they are both morphologically and

[6] This underlying sense of relativity is inherent, by definition, in any analysis that deals with standardized distances from a sample mean.

phonologically simple (e.g. Koiari, Yoruba, and Amele); seven languages (~21.9%) (e.g. Ju|'hoan, Nandi, and Vietnamese) are morphologically simple but phonologically complex; seven languages (~21.9%) are both morphologically complex and phonologically complex (e.g. Wichita, Lakhota, and Cairene Arabic); and seven (~21.9%) are morphologically complex but phonologically simple (e.g. Koasati, Yimas, and Ngarinjin). Languages in the upper-left and lower-right quadrants exemplify the negative correlation hypothesis while those in the upper-right and lower-left contradict it.

Table 5 reports the results of the Pearson product moment correlation (r), using the inflectional synthesis (IS) score, the log-transform potential syllable (Log-σ_p) score, along with the counts of consonants, vowels, and syllable types (σ) as variables. Significance (two-tailed) is reported by using (*) where ($p<0.01$).

TABLE 5. Correlation coefficients (r) for inflectional synthesis (IS), log transform of potential syllables (Log-σ_p), phonemic consonants (C), phonemic vowels (V), and syllable types (σ). (*) indicates significance ($p<0.01$).

	IS	Log-σ_p	C	V
Log-σ_p	0.070357			
C	-0.3226	0.32101		
V	-0.30912	0.51611*	0.62868*	
σ	0.37774	0.47023*	-0.17148	-0.0411

The results of the correlation test can be summarized in four statements: (1) Inflectional synthesis (IS) correlates positively with phonological complexity (Log-σ_p) but the correlation is not significant below the 0.01 level. Thus (1) can be taken neither as support for the negative correlation hypothesis nor for a positive correlation hypothesis. (2) IS correlates positively with number of syllable types (σ), though the correlation, again, is not significant at the 0.01 level. (3) The number of consonants in the phonemic inventories of the sampled languages correlates positively (and significantly) with the number of vowels but this is of no particular consequence to the negative correlation hypothesis. (4) Phonological complexity (Log-σ_p) correlates positively and significantly (at the 0.01 level) with number of vowels, suggesting the important role of vowels in determining the potential number of syllables. However, since V is not correlated significantly with Log-σ_p, the result is not germane to the negative correlation hypothesis.

4.3. Resampling results

Due to the admittedly small size of the sample and the anticipated influence of outliers, a resampling procedure known as the bootstrap was performed. Languages were randomly sampled with replacement[7] 1,000 times and the correlation coefficients (r) between IS and Log-σ_p were computed each time, generating the one thousand correlation coefficients shown in the histogram of Figure 7. The number of elements in each bootstrap sample equals the number of elements in the original data set. The range of correlation coefficients between IS and Log-σ_p obtained in this manner helps establish the uncertainty of the original estimated correlation between IS and Log-σ_p ($r=0.0704$).

As Figure 7 illustrates, the range of coefficient estimates among the bootstrap samples varies from about -0.5 to 0.6, distributed approximately normally about a mean of 0.061. A correlation coefficient is about as likely to be negative as it is to be positive, lending little credibility to a correlation (whether negative or positive) between IS and Log-σ_p. This confirms the result in §4.2, where it was shown that the correlation coefficient ($r=0.0704$) failed to achieve significance at even the 0.05 level ($p=0.702$).

[7] Every language was returned to the data set after being sampled. Thus, a given language from the original 32-language set could appear multiple times, randomly, in a given bootstrap sample. Each bootstrap sample consisted of 32 languages so if one language appeared multiply in a given bootstrap sample then another language was of necessity excluded. Thus, the procedure randomly includes or excludes outliers (to a randomly variable degree) in each bootstrap sample.

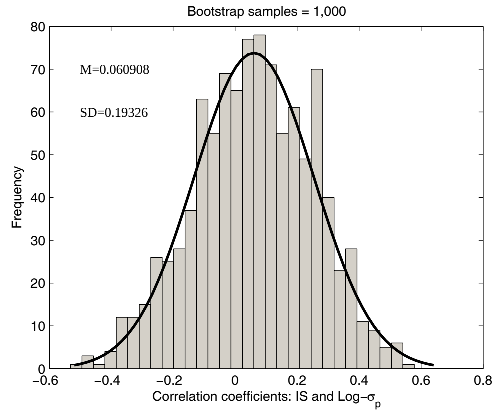

FIGURE 7. Frequencies of Pearson product moment correlation coefficients (r) using the bootstrap. The resampled correlation coefficients have an approximately normal distribution (M=0.061, SD=0.19). Furthermore, the estimates lie approximately on the interval [-0.5, 0.6]. These facts suggest that the correlation coefficients are just about as likely to be negative as positive. Results of this technique suggest that IS and Log-σ_p are neither positively nor negatively correlated.

5.0. Discussion

It is widely held that a "language which is simple and regular in one respect is likely to be complex and confusing in others. There seems to be a trading relationship between the different parts of the grammar which we do not fully understand" (Aitchison 1991: 214). In the present study, this "trading relationship" has been cast in terms of a scientific hypothesis and evaluated statistically. The data from the present 32-language sample provide no significant evidence against the null hypothesis of uncorrelated complexity. In other words, there is presently no scientific basis for claiming that a correlation exists. I conclude that the dictum, "All languages are equally complex," is dogmatic. Statements of this sort should be used with greater caution—if not discarded altogether—until such time as falsifiable, quantitative evidence of correlated complexity is brought forward.

While there are certainly many languages with simple phonology and complex morphology (or vice versa), there are also plenty of counterexamples—languages that are complicated (or simple) in both their phonological and morphological domains. While proponents of negative correlation may understandably cite Ju|'hoan and Vietnamese to bolster their argument, they must not ignore Koiari and Wichita, which present us with a much different result (see Figure 6). My findings represent a partial answer to a question posed by holistic typologists for generations (cf. Plank 1998) and should sober any who are tempted to make sweeping generalizations about the complexity inherent in human language.

Nevertheless, I am aware of the limitations of the present study and cannot therefore sweep out the negative correlation hypothesis entirely. As noted earlier, I have consciously chosen to test the negative correlation hypothesis only as it applies to inflectional morphology and certain aspects of phonology. It is entirely possible that, for example, languages make trade-offs between lexemes and instances of syntactic embedding, or even between derivational morphology and phonological "regulations." Tests of such hypotheses must await more ingenious metrics of complexity than are presently contemplated. Hopefully, the foregoing contribution, along with recent publications on complexity by Dahl (2004), Kusters (2003), and Trudgill (2004), will stimulate more thought on how to accurately sample it.

Now, a word about falsification: A scientific hypothesis is falsifiable to the extent that there exists a logically possible observation that is inconsistent with the hypothesis (Popper 1959). The hypothesis that inflectional morphology and phonology are uncorrelated is falsifiable. It can be falsified by observing any number of genetically-diverse, well-described languages with high phonological and low morphological complexity, and vice versa (using the same measurements discussed here). The more languages that are determined to have these characteristics, the weaker the hypothesis that inflectional morphology and phonology (as I have measured them) are uncorrelated. Of course, the engineering of new metrics which produce contradictory results could also falsify the hypothesis as long as such metrics are judged to be comparable.

There are various kinds of data, however, that do not falsify the hypothesis. For example, languages not included in my sample that are shown to have extremely high phonological complexity do not constitute falsification unless they are coupled with low inflectional synthesis scores. Also, data from languages for which phonologists have difficulty determining the exact number of phonemes do not falsify the hypothesis, since I have deliberately eschewed languages for which sufficient data do not exist.

Kusters (2003: 12) goes so far as to conclude that the negative correlation hypothesis (which he calls "equi-complexity") is not falsifiable at all. He observes that "for every change in complexity it can be argued that there is another component, in another domain of language structure, pragmatics, or even culture, where the amount of complexity would be leveled out. There is no feasible case where this a priori statement can be shown to be false." In the history of science, falsified theories serve as the groundwork for future theories that better account for phenomena (e.g. the motion of objects was explained with increasing explanatory power by Aristotle, Galileo, Newton, and Einstein, each of whom falsified the theory of his predecessor(s)). Non-falsifiable theories represent an ideological "closed circle" in which scientific discovery has no place. I believe the only justifiable course in linguistic science is to submit hypotheses whose results can be either replicated or falsified. Naturally, mounting replications must yield acceptance and mounting falsifications must yield rejection.

6.0. Conclusion

My research is a preliminary step in assessing the negative correlation hypothesis. It suggests no significant relationship between the number of potential syllables and the number of verbal inflectional markers among the languages of the world. Skeptics may object that no one has ever claimed otherwise, at least not for this particular grammatical relationship. Nonetheless, if negative correlation is indeed a feature of human language, as implied by the widely embraced notion of grammatical equality, then surely there ought to be quantitative evidence of the "trading relationship" somewhere. Those who hold to the notion of grammatical equality should at the very least recognize that it has never been proven using any comprehensive, quantitative method.

At present, I have not found this trading relationship among potential syllables and verbal inflectional markers. As far as I know, no one else has tried to document correlated complexity across a diverse language sample. So, I am the first to admit that my findings do not sound a death knell for negative grammatical correlation—such is the nature of a preliminary, falsifiable hypothesis. For this reason, the findings deserve to be replicated or falsified using other language samples and other quantitative metrics of complexity. For now, the results of my research lead to the only possible conclusion at present: if there are indeed cognitive limits that determine the complexity of the components of a language, these limits are not, so far as we know, approached by existing languages.

Acknowledgments

I wish to thank Johanna Nichols and Ian Maddieson for providing the inspiration to carry out this research, and for the time they have devoted to commenting on various drafts. Three anonymous reviewers and John McWhorter provided a wealth of helpful insights and suggestions. David Rood and Juliette Blevins were kind enough to consult with me on the Wichita and Yurok data, respectively. Despite their various contributions, however, none of these people is in any way responsible for errors or omissions in the paper. This research was partially funded by a Jacob K. Javits Fellowship to the author.

References

Aitchison, Jean. (1991). *Language Change: Process or Decay?* Cambridge: Cambridge University Press.

Akmajian, Adrian, Richard A. Demers, Ann K. Farmer & Robert M. Harnish (1997). *Linguistics: An Introduction to Language and Communication*. 4th edn. Cambridge: MIT Press.

Arends, Jacques (2001). Simple grammars, complex languages. *Linguistic Typology* 5, 180–182.

Austin, Peter (1981). *A Grammar of Diyari, South Australia*. Cambridge: Cambridge University Press.

Awobuluyi, Oladele (1978). *Essentials of Yoruba Grammar*. Ibadan: Oxford University Press Nigeria.

Bickel, Balthasar & Johanna Nichols (2004). The AUTOTYP network of typological databases. http://www.uni-leipzig.de/~autotyp/.

Bickel, Balthasar & Johanna Nichols (2005). Inflectional synthesis of the verb. In Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.), *The World Atlas of Language Structures*, 94–97. Oxford: Oxford University Press.
Blevins, Juliette (2003). One case of contrast evolution in the Yurok vowel system. *International Journal of American Linguistics* 69, 135–50.
Blevins, Juliette (in preparation). Yurok Reference Grammar. Ms., University of California, Berkeley and Max Planck Institute for Evolutionary Anthropology, Leipzig.
Bybee, Joan (1985). *Morphology: A Study of the Relation between Meaning and Form*. Philadelphia: John Benjamins.
Bybee, Joan, Revere Perkins & William Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
Carter, Richard T. (1974). Teton Dakota Phonology. Doctoral dissertation, University of New Mexico.
Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper and Row.
Cipollone, Nick, Steven H. Keiser & Shravan Vasishth (1998). *Language Files: Materials for an Introduction to Language and Linguistics*. 7th edn. Columbus: Ohio State University Press.
Coate, H.H. & Lynette Oates (1970). *A grammar of Ngarinjin, Western Australia*. Australian Aboriginal Studies 25. Linguistic series 10. Canberra: Australian Institute of Aboriginal Studies.
Comrie, Bernard (1989). *Language Universals and Linguistic Typology*. 2nd edn. Chicago: University of Chicago.
Comrie, Bernard (1997). Turkish phonology. In A. S. Kaye (ed.), *Phonologies of Asia and Africa*, 883–898. Winona Lake, IN: Eisenbrauns.
Craig, Collette G. (1977). *The Structure of Jacaltec*. Austin: University of Texas.
Creider, Chet A. & Jane Tapsubei Creider (1989). *A Grammar of Nandi*. Hamburg: Helmut Buske Verlag.
Dahl, Östen. (2004). *The Growth and Maintenance of Linguistic Complexity*. Philadelphia: John Benjamins.
Davies, John. (1981). *Kobon*. Lingua Descriptive Studies 3. Amsterdam: North Holland.
DeGraff, Michel (2001). On the origin of creoles. *Linguistic Typology* 5, 213–311.
Dickens, Patrick (1994). *English-Ju|'hoan, Ju|'hoan-English Dictionary*. Köln: Koppe, 1994.
Donohue, Mark (1999). *A Grammar of Tukang Besi*. Berlin: Mouton.
Dutton, Tom E. (1996). *Koiari*. München: Lincom Europa.
Dutton, Tom E. (2003). *A Dictionary of Koiari, Papua New Guinea, with Grammar Notes*. Canberra: The Australian National University.
Everett, Daniel L. & Barbara Kern (1997). *Wari'*. New York: Routledge.
Foley, William A. (1991). *The Yimas Language of New Guinea*. Stanford: Stanford University Press.
Greenberg, Joseph H., Charles A. Ferguson & Edith A. Moravcsik (eds.) (1978). *Universals of human language* (4 vols.). Stanford, CA: Stanford University Press.
Guirardello, Raquel (1999). A Reference Grammar of Trumai. Doctoral dissertation, Rice University.
Hagège, Claude (2001). Creoles and the notion of simplicity in human languages. *Linguistic Typology* 5, 167–175.
Harrell, Richard S. (1957). *The Phonology of Colloquial Egyptian Arabic*. New York: American Council of Learned Societies.
Hartwig, F. & B.E. Dearing (1979). *Exploratory Data Analysis*. Newberry Park, CA: Sage Publications, Inc.
Holton, David, Peter Mackridge & Irene Philippaki-Warburton (1997). *Greek: A Comprehensive Grammar of the Modern Language*. New York: Routledge.
Kimball, Geoffrey D. (1991). *Koasati Grammar*. Lincoln: University of Nebraska.
Kuipers, Aert H. (1967). *The Squamish Language: Grammar, Texts, Dictionary*. Janua Linguarum 73. Paris: Mouton.
Kullmann, Rita (1996). *Mongolian Grammar*. Hong Kong: Jensco.
Kusters, Wouter (2003). *Linguistic Complexity: The Influence of Social Change on Verbal Inflection*. Utrecht: Landelijke Onderzoekschool Taalwetenschap (LOT).
Kusters, Wouter & Pieter Muysken (2001). The complexities of arguing about complexity. *Linguistic Typology* 5, 182–185.

LeFebvre, Claire. (2001). What you see is not always what you get: Apparent simplicity and hidden complexity in creole languages. *Linguistic Typology* 5, 186–213.
Miller-Ockhuizen, Amanda. (2003). *The Phonetics and Phonology of Gutturals: A Case Study from Ju|'hoansi*. New York: Routledge.
Maddieson, Ian (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.
McWhorter, John (2001). The world's simplest grammars are creole grammars. *Linguistic Typology* 5, 125–166.
McWhorter, John (2001a). What people ask David Gil and why: Rejoinder to the replies. *Linguistic Typology* 5, 388–413.
McWhorter, John (2005). *Defining Creole*. New York: Oxford University Press.
Meillet, Antoine (1906). Leçon d'ouverture au cours de grammaire comparée au Collège de France. Reprinted in Antoine Meillet (1921), *Linguistique historique et linguistique générale 1*. Paris: Klincksiek.
Mortensen, David (2004). Preliminaries to Mong Leng (Hmong Njua) phonology. Unpublished manuscript, University of California, Berkeley.
Newman, Paul (2000). *Hausa: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
O'Grady, William, John Archibald, Mark Aronoff & Janie Rees-Miller (2005). *Contemporary Linguistics: An Introduction.* 5th edn. New York: St. Martin's Press.
O'Grady, William, Michael Dobrovsky & Mark Aronoff (1997). *Contemporary Linguistics: An Introduction*. 3rd edn. New York: St. Martin's Press.
Okell, John (1969). *A Reference Grammar of Colloquial Burmese*. Oxford: London.
Payne, Thomas E. (1997). *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Oxford University Press.
Plank, Frans (1998). The co-variation of phonology with morphology and syntax: A hopeful history. *Language Typology* 2, 195–230.
Popper, Karl (1959). *The Logic of Scientific Discovery*. New York: Basic Books.
Rice, Keren (1989). *A Grammar of Slave*. New York: Mouton.
Rijkhoff, Jan & Dik Bakker (1998). Language sampling. *Linguistic Typology* 2, 263–314.
Roberts, John R. (1987). *Amele*. New York: Croom Helm.
Rood, David S. (1976). *Wichita Grammar*. New York: Garland.
Rood, David S. (1996). Sketch of Wichita, a Caddoan language. In Ives Goddard (ed.), *The Handbook of North American Indians* 17, 580–608. Washington: Smithsonian Institution.
Saeed, John (1999). *Somali*. Philadelphia: John Benjamins.
Seuren, Pieter A.M. (2001). Simple and transparent. *Linguistic Typology* 5, 175–180.
Smeets, Ineke (1989). Mapuche Grammar. Doctoral dissertation, University of Leiden.
Sohn, Ho-Min (1999). *The Korean Language*. Oxford: Cambridge University Press.
Spencer, Andrew (1998). *Morphological Theory*. Cambridge: Blackwell.
Sridhar, S.N. (1990). *Kannada*. London: Routledge.
Thompson, Laurence C. (1987). *A Vietnamese Reference Grammar*. Honolulu: University of Hawaii Press.
Trubetzkoy, Nikolaj. (1931). Die phonologischen Systeme. *Travaux du Cercle Linguistique de Prague* 4: 96–116.
Trudgill, Peter (2001). Contact and simplification. *Linguistic Typology* 5, 371–374.
Trudgill, Peter (2004). Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology* 8, 305–320.
Watkins, Laurel J. (1984). *A Grammar of Kiowa*. Lincoln: University of Nebraska.
Watson, Janet C.E. (2002). *The Phonology and Morphology of Arabic*. Oxford: Oxford University Press.
Whaley, Lindsay J. (1997). *Introduction to Typology: The Unity and Diversity of Language*. London: Sage.
Wurzel, Wolfgang Ullrich (2001). Creoles, complexity, and linguistic change. *Linguistic Typology* 5, 377–378.