VERSITA OPEN

SEÁN ROBERTS, JAMES WINTERS
University of Edinburgh

SOCIAL STRUCTURE AND LANGUAGE STRUCTURE: THE NEW NOMOTHETIC APPROACH

Recent studies have taken advantage of newly available, large-scale, cross-linguistic data and new statistical techniques to look at the relationship between language structure and social structure. These 'nomothetic' approaches contrast with more traditional approaches and a tension is observed between proponents of each method. We review some nomothetic studies and point out some challenges that must be overcome. However, we argue that nomothetic approaches can contribute to our understanding of the links between social structure and language structure if they address these challenges and are taken as part of a body of mutually supporting evidence. Nomothetic studies are a powerful tool for generating hypotheses that can go on to be corroborated and tested with experimental and theoretical approaches. These studies are highlighting the effect of interaction on language.

Key words: nomothetic, social structure, complex adaptive systems, linguistic niche hypothesis, cultural evolution

1. Introduction

A century ago, Edward Sapir discussed the relationship between language, human society and the environment:

Properly speaking, environment can act directly only on an individual, and in those cases where we find that a purely environmental influence is responsible for a communal trait, this common trait must be interpreted as a summation of distinct processes of environmental influences on individuals. Such, however, is obviously not the typical form in which we find the forces of environment at

Address for correspondence: Seán Roberts, Language Evolution and Computation Research Unit, School of Philosophy, Psychology and Language Sciences, The University of Edinburgh Dugald Stewart Building, 3 Charles Street Edinburgh, EH8 9AD, e-mail: S.G.Roberts@sms.ed.ac.uk; James Winters; e-mail: J.R.Winters@sms.ed.ac.uk

work on human groups. ... The important point remains that in actual society even the simplest environmental influence is either supported or transformed by social forces.

(Sapir, 1912, p. 226)

Sapir recognised that human culture at the society level and the environmental forces working at the individual level could co-evolve. This co-evolution may have begun before the evolution of modern language (e.g. Dunbar, 1993; Isbell and Young, 1996). Language, then, is a complex adaptive system: it emerges as a product of its underlying speech community, but also adapts to the very dynamics from which it emerged (Beckner et al., 2009). Recent research has considered how language structure and social structure might be linked based on newly available data and new statistical techniques. This approach can be broadly categorised as 'nomothetic': seeking law-like explanations for general phenomena from large sets of cross-linguistic data. On the other hand, more traditional research, categorisable as 'idiographic', uses careful case-studies to motivate hypotheses. Given this difference, there is a tension between proponents of the two approaches.

This paper will review some of the new nomothetic studies and try to place them in relation to other approaches. First, the differences between the nomothetic and idiographic approaches are illustrated with a short example. The next section gives an example of some nomothetic research. Section 4 discusses some problems with the nomothetic approach such as the possibility of spurious correlations. The next few sections consider how the nomothetic approach can be combined with other approaches such as experiments and computational modelling. These sections use case-studies to illustrate their points. Section 5 shows that nomothetic studies can be good for generating hypotheses. However, section 6 demonstrates that nomothetic results can be combined with other approaches, for instance in order to address the effects of social structures on how languages change. Section 7 shows that nomothetic approaches can provide insights for theories which other approaches would find difficult to generate, with an example from the literature on the critical period effect for language acquisition. Section 8 demonstrates that results from nomothetic studies can be a catalyst for bringing together researchers and techniques from different areas and focussing their attention on specific problems. Finally, we argue that results from experimental and modelling work suggest that phylogenetic techniques are the best methods for carrying out nomothetic studies.

2. What is the nomothetic approach?

The nomothetic and idiographic approaches have been used to study the co-evolution of language structure and social structure. These two approaches will be briefly illustrated with an example from studies of colour words. First, the idiographic approach: Given the co-evolutionary nature of language and society, the

lexicon of a society may reflect the values and activities that it considers relevant. As an extreme example, MacKeigan and Muth (2006) report that speakers of Tzotzil have hundreds of compound forms for describing various dimensions of colour. This is because the main industry of the society is hand-weaving colourful textiles. Weavers work in groups and need fine-grained labels with which to request certain threads. MacKeigan and Muth were able to track the spread of individual colour words within the social network of the weavers. In this way, the affects of social structure on the lexicon could be identified.

This is an example of the idiographic approach: A careful study of individuals in a particular society. Although this is a powerful research method, it is also complex, costly and time-consuming. Given the availability of new electronic resources shared over the internet, it is now possible to address the same topics simply, cheaply and quickly.

To demonstrate this point, we conducted a nomothetic analysis of the effect of colour terms on cultural aspects of society. Our hypothesis was that the range of vocabulary a society has with which to discuss colours should affect the range of colours in artifacts whose appearance has been negotiated, such as national flags.

Indeed, a multiple regression shows that the number of basic colour categories in a language is a significant predictor of the number of colours in the national flag for speakers of that language (r = 0.12, F(16,106) = 1.86, p = 0.03, controlling for the effects of speaker population, distance from the equator – see Bornstein, 1973 – and a country's per-capita gross domestic product as a measure of industrialisation – see Kay and Regier, 2003. Data from Kay and Maffi, 2008; Wikipedia, 2011; International Monetary Fund. Online, 2011). This appears to support the hypothesis that language structure can affect cultural artifacts.

This statistic is an example of a nomothetic approach – considering data across many languages and sources (although admittedly using crude statistical techniques). It was done using an ordinary laptop, took less than a day to complete and the data and analysis cost nothing to obtain. The advantages and attraction of such an approach are clear. However, there are many problems – some of which will be discussed in this paper. These include problems with the data and statistical analyses, the assumptions behind the analyses and the explanatory power of the analyses.

Given that hypotheses from the idiographic approach come out of detailed studies of causal processes while nomothetic studies look at correlations between large-scale variables, it is not surprising that there is a tension between proponents of each approach. Although, as we will argue, there are many problems with the nomothetic approach, a common source of this tension is often a misunderstanding about the role of the nomothetic approach. Many antagonists argue that the nomothetic studies are poorly motivated and do not provide support for the theories of interest. However, many researchers using nomothetic approaches see its role as exploratory and to demonstrate features and interactions of interest that other approaches may test, not necessarily a definitive proof in itself. The

candid introduction of Hay and Bauer (2007) demonstrates that researchers are aware of this:

> To our knowledge, no-one has ever reported a statistical correlation between the number of speakers of a language, and how many phonemes that language has. This, of course, is not surprising: why would anyone look for such an association in the first place? It is certainly not an association one would necessarily expect. However in the process of proof-reading a manuscript that contained information about series of languages, including their populations and vowel inventories, it struck us that there appeared to be some link. We couldn't resist checking this apparent link more systematically. In this short report we show that there is, indeed, an association between phoneme inventory and population size. We do not have well-developed theoretical arguments to offer about why this should be. However the correlation seemed intriguing enough that it was worth simply publishing the result, and leaving it up to readers to draw their own conclusions.
>
> (Hay and Bauer, 2007, p. 388)

The tension between proponents of the two approaches will be discussed throughout the paper but we argue that both approaches feed into each other: Nomothetic approaches can help generate hypotheses to be tested, but other approaches are needed to test them. We continue with a short review of some nomothetic studies.

3. Nomothetic approach to language

Recently, the amount of data available on language typologies and social features has increased due to database technology. These data, along with tools and techniques with which to analyse them, have become more accessible through the internet. This has prompted large-scale statistical investigations.

For example, Lupyan and Dale (2010) show that a community's population is correlated with the level of morphological complexity in their language (morphological complexity includes variables that measure the complexity of the case system and is contrasted with lexical strategies, for example marking the future tense morphologically as opposed to lexically). Language typology data from the World Atlas of Language Structures (WALS, Haspelmath et al., 2008) is combined with data on the speaker population and geographic spread from the Ethnologue (Gordon, 2005). A general trend is identified: Languages with a small number of speakers, low geographic spread and few linguistic neighbours tend to have high morphological complexity (esoteric languages).

On the other hand, languages with a large number of speakers, large geographic spread and many linguistic neighbours tend to have low morphological complexity (exoteric languages).

The Linguistic Niche Hypothesis (LNH) put forward by Lupyan & Dale proposes that, just as the structure of biological organisms are affected by their ecological niche, the structure of society affects the evolutionary pressures on language structures. In particular, a distinction is made between exoteric and esoteric niches. An exoteric niche contains languages with a large variety of speakers putting pressure on the communal language to become suited for communication between strangers and become more learnable by non-native (adult) learners (Lupyan & Dale, 2010). English, Swahili and Hindi are all examples of languages emerging from exoteric niches, in that they are more likely to be learned or taught by non-native speakers and be used with individuals from different cultural backgrounds (ibid). Conversely, esoteric niches are composed of languages like Tatar, Elfdalian and Algonquin (ibid). All of these languages belong to relatively small populations where individuals are part of a tightly knit community, based on a shared cultural and social identity.

Given these differing linguistic niches, Lupyan & Dale's main proposal is that the morphological features of a language are the product of adaptation to learning constraints and the communicative requirements of the underlying speaker population. Therefore, a language being subjected to a greater number of outside learners – for instance, via colonisation or large-scale migration – is thus under a greater pressure to become learnable, subsequently simplifying its morphology and increasing the productivity of existing grammatical patterns.

Proponents of the idiographic approach also argue that the amount of language contact has an effect on language complexity (Trudgill, 2004; DeGraff, 2001; Dahl, 2004), although opinions differ on the importance of this factor. Lupyan and Dale cautiously claim that the relationship is only partial, and Trudgill (2004) argues that population size, network structure and language contact need to be considered together. On the other hand, McWhorter (2008) suggests that the level of non-native language acquisition is the only factor that affects the simplification of grammar over generations. Other theories highlight the importance of background knowledge shared by a small community (Nettle, 1999; Wray and Grace, 2007). Recent experimental evidence also supports the hypothesis that second language learners find morphology more difficult to master than native speakers (Clahsen et al., 2010). It's worth noting that nomothetic approaches do not always contradict conclusions from other approaches, and are not always the most extreme.

Languages in exoteric niches also become more analytical, which, according to Wray and Grace (2007), increases their form-meaning compositionality in that "meanings of expressions can be determined from their composition, because the system approximates a one-to-one relationship between forms and meanings" (pg. 9). On the basis of similar claims, Carstairs-McCarthy (2005) suggests that linguistic universals are an emergent property of exoteric language use over many generations, rather than being biologically based.

4. Challenges for the Nomothetic approach

The nomothetic approach can be a powerful tool for exploring relationships between linguistic and social variables. However, the approach can also be problematic. In Roberts and Winters (2012), we discuss these problems and summarise them below.

Nomothetic approaches use large-scale, cross-cultural datasets. Since the inference in nomothetic approaches derives directly from the data, these studies are sensitive to the quality of the data. This may be especially difficult to verify if there is a large amount of data on a wide variety of languages or cultural practices, since any one researcher would be unlikely to have in-depth knowledge about each. Furthermore, when integrating information on different languages into a single typological framework, the assumptions of the framework can be imposed on the data. There is also the issue of temporal resolution. Many nomothetic approaches suggest that linguistic and cultural variables co-evolve with each other over time, but the data that is used to test this is often from only one period in time.

Perhaps a bigger danger is the possibility of spurious correlations. Large datasets with multiple variables are likely to contain correlations between variables that have no direct causal link (see Roberts and Winters, 2012). For example, countries in which the acacia tree Acacia nilotica grows are significantly more likely to have tonal languages (χ^2 with Yates' continuity correction = 47.1, df = 1, p < 0.0001, data from Crop Protection Compendium, 2008 and Maddieson, 2008). However, a causal link is, obviously, very unlikely.

The more problematic cases are where causal links may be more intuitive. For example, a colleague challenged us to find a relationship between two variables that they thought could not possibly be linked: basic word order and the average number of offspring that parents have. However, a robust correlation was apparent. Data was collected from the World values database (World Values Survey Association, 2009) and the World Atlas of Language Structures (Dryer, 2011). The following variables were entered into a linear regression: basic word order of language, age group, sex, employment status, marriage status, level of education and religion. Controlling for the other factors, basic word order was a significant predictor of the number of children a couple have ($F(3,35030) = 121.6$, p < 0.0000001). This model accounted for 36% of the variance in number of children. People speaking SOV languages were likely to have more children than those speaking SVO languages.

As Hay and Bauer (2007) note, these kinds of results are intriguing, and it is tempting to create stories that would fit the results. Could this result reflect a similar mechanism to the one suggested by Lupyan and Dale (2010)? That is, a greater number of children in the population affects the linguistic 'niche' that languages must adapt to, therefore influencing the way the languages change over time. However, the danger is that a large enough data set can provide apparent

support for any hypothesis. There are two basic problems. To begin, the strength of the correlation may be impressive, but it must be considered alongside alternative hypotheses. For instance, whether participants speak a language with front rounded vowels is a better predictor of the number of children an individual has ($F(1,32439) = 533.9$, $p < 0.0000001$). In fact, calculating the correlation between the number of children and all variables in the World Atlas of Language Structures, word order ranks 11th (within top 15%). This suggests that the result is a spurious correlation that is brought about either because of the relatedness of languages and cultures or because of complex relationships with other variables.

Secondly, the correlation does not suggest the direction of causation. Nomothetic studies can demonstrate that variables are statistically linked, but can't determine the mechanism which brought about the correlation. That is, on their own, nomothetic studies only have weak explanatory power (see Roberts and Winters, 2012).

As an example, Lupyan and Dale (2010) suggest that the number of second language learners is a factor in the explanation of the distribution of morphological complexity. Although it is generally true that ultimate attainment in language acquisition is correlated with age of first exposure, the cause of this is difference is contested. This makes it difficult to identify the causal factor driving the relationship between second language learners and morphological complexity. For instance, the linguistic output of a second language learner (and so the input to the next generation) may be affected by social aspects: At the individual level, second language learners may avoid constructions that they find difficult to avoid embarrassment (e.g. Schachter, 1974). Output may also be affected by lack of motivation, low self-esteem or anxiety (Krashen, 1982), which may differ in children and adults and may result from other social factors.

Furthermore, adults and children differ in the way they interact with others in their society in terms of the number of interlocutors, frequency, context and function of the interaction (Kuhl et al., 1997). Also, second language learners may have a high functional attainment in the specific domains involved in cross-linguistic contact (Hyltenstam and Abrahamsson, 2003; Bialystok and Miller, 1999) and age of acquisition may affect different aspects of language to different degrees (e.g. Eubank and Gregg, 1999).

At the society level, ultimate attainment is affected by the linguistic distance between the first and second language (e.g. Lado, 1957; Kellerman and Sharwood Smith, 1986). This means that it's not just the number of other languages a community has contact with that should affect its evolution, but the properties of those languages. The picture is complicated when considering that social groups split and merge and borrow linguistic features from one another. For example, Matisoff (1978) sees the relations between languages as "an interlocking network of fuzzy-edged clots of languages, emmitting waves of mutual influence" (Matisoff, 1978, p. 2) rather than neat tree-like dispersion. Furthermore, borrowing from other languages may not necessarily cause a systematic change in the borrowing language (e.g. English

has borrowed the word 'taboo' from Tongan but the influence is relatively minute in shaping the historical trajectory of English, Greenhill et al., 2009).

In summary, although the proportion of second language speakers in a community is related to the morphological complexity of its language, the mechanism that brings the relationship about may be complex. However, while there may be problems with the nomothetic approach, we suggest that it can still be a productive tool for research when combined with other approaches. Indeed, Lupyan and Dale claim that their results show a strong tendency which suggests a hypothesis that can go on to be tested directly by other approaches.

5. Generating hypotheses

The increasing availably of data and sophistication of statistical techniques means that generating hypotheses with a nomothetic approach is increasingly easy. While there are concerns about the statistical methods, the nomothetic approach of Lupyan and Dale (2010) does generate a testable hypothesis: the balance in cognitive profiles within a speech community will affect the selection pressures on language and, over generations, the language will change to accommodate its hosts. This view of language evolution is compatible with the view of language as an organism that adapts to the cognitive niche of human brains (e.g. Christiansen and Chater, 2008; Smith and Kirby, 2008). However, before the hypothesis can be tested, a specific mechanism with predictive power should be identified. Without one, the hypothesis is difficult to falsify.

As an example, the differences between first and second language acquisition performance have been linked to different use of procedural and declarative memory (Ullman, 2005; Paradis, 2004; Digiulio et al., 1994). Procedural memory is linked with automatic processes such as morphological and syntactic operations and declarative memory is linked with conscious retrieval such as retrieval of lexical items. The ability to incorporate knowledge into procedural memory atrophies over time. Second language learners, therefore, should acquire lexical marking better than morphological marking.

This predicts that any stratification of the procedural/declarative memory capacity of a population, in addition to child versus adult learners, will affect the morphological complexity of its language. For instance, procedural/declarative profiles are affected by gender (Hartshorne and Ullman, 2006), daytime napping (Backhaus and Junghanns, 2006) and alcohol (Smith and Smith, 2003). One might be able to show that morphological complexity varies with male to female ratio, sleeping patterns (e.g. siestas) and per-capita alcohol consumption. A preliminary investigation found that the number of grammatical categories in a language is correlated with the sex ratio of under 15s of its country (134 languages, $r = 0.17$, $p < 0.05$, data from Bickel and Nichols, 2008, Central Intelligence Agency, 2009). However, this correlation goes in the opposite direction to that predicted by the

above hypothesis (populations with more males than females tend to have fewer grammatical categories).

While this result might appear to dismiss this particular mechanism as a candidate for Lupyan and Dale (2010)'s hypothesis, we suggest that this may not be the right approach. In seeking to test the hypothesis generated by looking at a large amount of low-quality data, more data of the same kind has been added. That is, a nomothetic hypothesis has been tested by a nomothetic study, which is vulnerable to the same kinds of weaknesses. Instead, hypotheses generated by nomothetic approaches should be tested using different approaches. An experimental approach might be more appropriate. Indeed, recent experimental work has focussed on the differences between how native and non-native learners adults acquire morphology (e.g. Clahsen et al., 2010). Alternatively, what is required is a way of combining large-scale, cross-cultural, society-level data with small-scale, individual-level data. New phylogenetic approaches (see section 9.1), could be the key to this integration.

6. Combining approaches

Nomothetic studies are suggesting that the structure of a society has an effect on the structure of the language that is used within it. However, the sections above demonstrated that nomothetic approaches have low explanatory power in isolation. This section considers how nomothetic results can be combined with experimental and modelling approaches to test and extend theories of the relationship between language structure and social structure.

The nomothetic approach is a powerful tool for generating testable hypotheses. Yet applying these tests to real-world situations is often complicated, costly and affected by factors outside the scope of the hypothesis. These principles can, however, be extended by integrating methodologies and hypotheses within a general theoretical framework. Complementing the nomothetic approach are models and experiments.

Both experiments and agent-based modelling are powerful tools for looking at how interactions between individuals can affect phenomena at the society level. The transmission of cultural features, such as language, is constrained by the social structure (Chambers, 1995). For instance, human societies involve interactions between individuals in so-called social networks (Newman and Park, 2003) which differ from other naturally occurring networks in many ways, including the 'small-world' effect. That is, your friends tend to be friends with each other and popular people tend to know other popular people (and vice-versa). Linguistic interactions, then, are constrained by social networks in that they are not merely achieved via random contacts (Beckner et al., 2009). A consequence of this is that social structure is predicted to pattern with linguistic variation (Weinreich et al., 1968) and act as the source and product of language change Croft (2000, 2010). It is through the interaction of individuals that novel linguistic variants emerge, and

are subsequently selected and integrated within a community though propagation and fixation.

A serious limitation of nomothetic and idiographic studies is their lack of temporal resolution: the long time scale of historical change precludes the linguist from performing a sufficient number of multiple observations at small but regular intervals (Fagyal et al., 2010, p. 3). Agent-based models provide one interesting route for exploring the role of social structure in the diffusion process. One recent application has been to test earlier theories put forward by Bloomfield (1933), Séguy (1973) and Trudgill (1974) (cf. Nerbonne, 2010). Indeed, Bloomfield realised that the amount of dialect differentiation caused by diffusion depends on the communication density (Nerbonne, 2010, p. 3822). Fagyal et al. (2010) used a series of simulations to show that innovations spread following an S-shaped curve and stabilise as norms only when the network is: (1) socially heterogeneous (scale-free), with tight-knit communities (high clustering) that keep their members in close reach (small diameter), and (2) group members imitate the language use of those whom all perceive as the most popular (ibid. p. 17). Additionally, the study highlighted some of complex dynamics at play, with the need for both highly-connected agents (leaders) and relatively isolated individuals (loners): here, leaders help establish the appearance of linguistic norms, whilst loners help maintain innovation through the introduction of novel variants.

Besides the well-studied dynamics of linguistic features in social networks (cf. Milroy and Milroy, 1985; Bogunã et al., 2004), the transmission of these variants is not often considered to follow a co-evolutionary relationship with social structure. Quillinan et al. (2010) present a co-evolutionary model of social structure and cultural learning. Social interactions between individuals are constrained by the social network, but the social network is also dynamically affected by the social interactions of individuals. The model contains a population of agents which each have a number of linguistic traits. The social distance between two agents is calculated from the difference between their linguistic features. This implements a homophily bias (McPherson et al., 2001): Agents tend to form social bonds based with agents who are similar to them. The structure of the social network evolves by three principles: Forming bonds with similar agents, breaking bonds with dissimilar agents and learning between bonded agents. Over time, social networks emerge with properties similar to human social networks. The key to this emergence is the learned aspect of the cultural transmission. The rate of learning affects both the social structure and the type and distribution of the languages.

Considering population density and other demographic variables chimes with evolutionary models in biology (cf. Hawks, 2008), culture (Shennan, 2000; Henrich, 2004; Richerson et al., 2009; Lycett and Norton, 2010) and linguistics (Nettle, 1999; Wichmann and Holman, 2009; Vogt, 2009). In tandem with these general models are several recent papers looking a demography and phonemic diversity (de Boer, 2000; Trudgill, 2004; Pericliev, 2004; Hay and Bauer, 2007; Winters, 2010; Atkinson,

2011b). Here, we have seen the integration of several theoretical and methodological approaches. For instance, when investigating the emergence of vowel systems in agent-based computer simulations, de Boer (2000) manipulated the size of the population to find that "the success of all population sizes is comparable, but the vowel system size of small populations is smaller than that of large ones, reflecting the lower stability" (pg. 459)This has allowed for the development of theoretical work (Trudgill, 2004) and the application of statistical methods to real-world phoneme inventory data (Hay and Bauer, 2007). The general findings all point toward a relationship between the phoneme inventory size and the number of speakers, yet there is much more work needed to determine whether this relationship is causal.

Ferdinand and Zuidema (2009) also show that social structure should affect a rational agent's approach to the learning problem in general. In an iterated learning model with Bayesian agents, language learning is modeled with generations of agents estimating the 'language' of the previous generation. Agents integrate the data that they hear and their own prior bias for which languages are more likely (both innate and learned biases). Over many generations, the distribution of 'languages' in the population will eventually converge to a value that reflects the prior bias of the individual. However, Ferdinand and Zuidema show that the population size can also affect the eventual distribution of the languages. Burkett and Griffiths (2010) also show that manipulating the individual's expectations about the structure of society can affect the distribution of languages that eventually emerge. Therefore, languages may evolve in different ways within different kinds of social structure.

However, Ferdinand and Zuidema also point out that the scope of a rational agent's biases is contentions. A fully rational agent should take the social structure into account when learning, not just the properties of the languages. It is an empirical question to what extent humans integrate the distribution of variation over speakers into the way they learn.

Laboratory experiments are also introducing innovative ways of merging the robust time-scales of agent-based modelling whilst generating real-world data. One particular source of success in recent years has been the testing and contrasting of different cultural transmission dynamics (cf. Mesoudi and Whiten, 2008; Scott-Phillips and Kirby, 2010). Garrod (2010), for instance, found that by contrasting two transmission dynamics – vertical cultural transmission and repeated pairwise interaction – chains of learners would produce different outputs. Specifically, the graphical communication systems that emerged were better suited to their specific forms of transmission: forms transmitted vertically tended to retain their iconic appearance, whereas the interactively produced forms decreased in complexity through taking on a more symbolic aspect. Tamariz et al. (2012) also used an iterated learning experiment with human participants to demonstrate that interaction with interlocutors, generation turnover and negotiation between interlocutors were crucial factors in determining the kinds of structure that emerged in the linguistic system.

Social structure is also shown to affect an individual's sensitivity to the distribution of variation across speakers. Roberts (2010) set up an experiment where players could only interact through an 'alien' language via an instant-messaging system, with the aim being to trade commodities in a series of rounds. The results showed that, if players interacted frequently enough with their team-mates and were in competition with another group, then linguistic diversity emerged. Over the course of the game each team developed its own 'variety', and this was used as a marker of group identity.

This demonstrates a relationship between linguistic diversity at the community level and individual identity (possibly explaining the relationships between phonetic diversity and group size, discussed above). Nettle (1999) discusses similar situations in Papua New Guinea, but based on co-operative rather than competitive principles: The richness of the land means that trading physical commodities is redundant. However, groups still need to co-operate for politics and war. Instead of investing materially, they invest culturally by learning the other communities' language. It is common for children to be sent to other communities to learn their language. There are even cases of communities consciously agreeing to introduce variation in order to distinguish themselves from other communities. Linguistic diversity, then, becomes a commodity which is maintained (see also Nettle and Dunbar, 1997). Crucially, as in Roberts (2010), diversity emerges and is maintained because the linguistic system has come to encode the speaker's identity within their society. That is, utterances refer to properties of the speaker as well as the objects and actions they pick out in the world. In this way, the function of a language may be radically altered by the structure of the society in which it is embedded. New experimental paradigms can help test these hypotheses. As the examples above demonstrate, theories about the relationship between social structure and language structure can be informed by combining evidence from multiple approaches.

7. Providing insights

While other approaches can test the hypotheses generated by nomothetic studies, principles suggested by nomothetic studies can also provide insights that can inform the research in other approaches. This is illustrated with an example from the work on the critical period effect for second language acquisition.

Experimental work has shown that age of exposure is a crucial factor for ultimate attainment of a language (e.g. Johnson and Newport, 1989, see Long, 1990; Birdsong and Molis, 2001 for reviews), although a definite cut-off age is disputed (see Hyltenstam and Abrahamsson, 2003).

However, Hurford (1991) argues that Johnson and Newport's (1989) results are also compatible with the hypothesis that the acquisition of a first language interferes with the acquisition of a second. This is, of course, a difficult confound to avoid with real participants, but can be overcome using computational modelling.

Hurford used a computational model to argue that a sensitive period for language learning at an early age should be expected. Individuals in the model had a set of genes that specified the period in their lives when they were most sensitive to acquiring language. If knowing a language for longer gives an individual a greater advantage, then individuals should evolve to be good at learning language early in their lifetime. Indeed, in the model, over evolutionary time, this genetic specification evolved so that individuals were most sensitive in early stages of their life. Rather than a diminished ability in adulthood, Hurford sees this as an adaptive 'boost' in acquisition ability during childhood.

However, the nomothetic studies above suggest that age is not the only factor relevant to this debate. Lupyan and Dale (2010) demonstrate that some languages may have adapted to the acquisition abilities of adults to a greater extent than others. This leads to a testable prediction: that ultimate attainment of large languages like English will be better on average than for small languages like Welsh, all other things being equal. That is, as well as humans adapting to when they are most sensitive to language acquisition, languages are adapting to the population of learners, meaning that there is a co-evolution between the two processes.

This effect would be difficult to spot using an experimental approach, since a lot of compatible studies with different languages would be required. It is also not an intuitive aspect to include in a computational model. While the nomothetic approach may not have explanatory power on its own, insights about the important factors of a debate from a nomothetic approach can feed back into both the modelling and experimental literature.

8. Nomothetic Approach as a Catalyst

Atkinson's (2011b) paper on phonemic diversity shows how a nomothetic approach can serve as a catalyst for bringing together researchers from disparate areas in linguistics. Based on previous work suggesting that the phoneme inventory size is related to socio-demographic dynamics (Hay and Bauer, 2007), Atkinson hypothesised phonemic diversity is concomitant with the geographic distance from Africa. He uses data from WALS and a series of multiple regressions to arrive upon two conclusions. First, the patterns of phonemic diversity across the globe supports a Serial Founder Effect (SFE) model, in which repeated population bottlenecks throughout human history gradually reduced the diversity. Second, in mirroring human genetic diversity, Atkinson argues these bottleneck effects on diversity point toward language as having putative origin in Africa.

Since the publication of these results, Atkinson's paper has drawn significant attention from various academic quarters. The paper elicited several responses in Science (cf. Atkinson, 2012), as well as a special issue of Linguistic Typology (cf. Bybee, 2011), that broadly challenged Atkinson's findings on both theoretical and methodological grounds.

One overarching theme from the theoretical side of the debate is challenging the links between phoneme inventory size, demography and cultural transmission. Peter Trudgill (2004; 2011) deserves a notable mention here for having repeatedly pointed out that these cross-linguistic similarities in phoneme inventory sizes are contingent on numerous social, cultural and demographic factors. Indeed, unrealistic assumptions can lead us to believe there is a linear relationship between a linguistic variable (e.g. phoneme inventories) and a socio-demographic variable (e.g. population size). The reality is likely to be far more complicated and not necessarily tractable for these coarse statistical methods, with any one factor potentially being misleading and not necessarily a reliable proxy variable (cf. Roberts and Winters, 2012).

Several critiques also emerged through utilising different methodological approaches. These ranged from the types of statistic tools employed (Jaeger et al., 2011), the sources of data (Wichmann et al., 2011; Cysouw et al., 2012) as well as different ways of quantifying phonemic diversity (Hunley et al., 2012). Hunley et al. (2012) offered one of the most prominent critiques in this respect: here, the authors compared a linguistic sample, consisting of 725 widespread languages with 908 distinct phonemes, to a genetic sample composed of 614 autosomal microsatellite loci in 100 widespread populations. By encoding the distribution of phonemes in an analogous manner to that of the genetic samples they found crucial differences in the way these two pattern. Whereas genetic patterns are consistent with predictions made by SFE, phonemic data violated these model assumptions. The rationale for this is largely because phonemes on average change more rapidly than genes and are therefore not stable enough to accurately inform us about possible ancient evolutionary traces.

The phonemic diversity debate should be held up as an example of how the nomothetic approach can serve as a tool for furthering theoretical and methodological contributions to the field. Despite the numerous disagreements that inevitably stem from these debates, the position taken here is that we should view this as a positive development. As Atkinson himself offered in a response:

> There are clearly interesting geographic patterns in global phonological (and typological) diversity that we are only beginning to understand. It seems like ancient demography and large-scale population migrations must be a leading contender for explaining these patterns... to build a more compelling case we need additional data, more evidence that a founder effect does indeed operate on phonemic diversity, and more effort to exclude alternative hypotheses. I agree with these general sentiments and hope that some of the discussion in this issue has helped move us in the right direction.
>
> (Atkinson, 2011a, p. 329-330)

Results from nomothetic studies can thus provide the focus for a debate between these various sub-disciplines, and offer insights into the theoretical and methodological differences between them.

9. Cultural Transmission Constrains Valid Nomothetic Methods

If we accept that language is not only a conveyer of cultural information, but is itself a socially learned and culturally transmitted system, then an individual's linguistic knowledge is the result of observing and repeating the linguistic behaviour of others (Kirby and Hurford, 2002). Attempts at capturing this process experimentally have found success under the rubric of Iterated Learning (henceforth, IL): a cycle of continued production and induction where individual learners are exposed to a set of data, which they must then reproduce and pass onto the next generation of learners (Kirby et al., 2008).

By taking into account the effects of cultural transmission, researchers can now probe for biases that were initially hard to spot due to their weak influence at the individual level (Kirby et al., 2007; Griffiths et al., 2010; Reali and Griffiths, 2009; Smith and Wonnacott, 2010). These studies demonstrate that weak biases may be amplified through cultural transmission meaning, that we cannot assume *a priori* that all universal patterns are the result of strong biases acting upon learning. This limits the explanatory power of the nomothetic approach.

The possibility of weak biases showing strong effects through cultural transmission means we must be careful when drawing conclusions from nomothetic data. Still, experimental methods can be complimentary as well as constraining, and we should therefore strive to use these as convergent processes of gathering evidence. Vogt (2009) provides a useful case study for combining computer modelling with real world data: here, he found that in models where communicative success is difficult, small populations show a tendency toward holistic systems, whereas larger populations are more inclined to compositional languages. This lends support to previous theoretical (Wray and Grace, 2007) and comparative (Levinson, 2006) results suggesting that languages spoken in larger communities tend to be more compositional.

9.1. Phylogenetics

Phylogenetic methods may provide a way of avoiding some of the difficulties mentioned above. This is a growing area of interest, and something that has clear methodological implications for testing language structure and social structure. Phylogenetic techniques in the field of evolutionary biology are a well-established, powerful set of tools that allow the testing of evolutionary hypotheses. More recently these methods are being applied to the analysis of linguistic and cultural data. In particular, the use of phylogenetics has led to the observation that languages evolve in punctuational bursts (Atkinson et al., 2008), tested multiple hypotheses for word order universals (Dunn et al., 2011), explored the role of population movements (Kaestle and Smith, 2001), examined the rise and fall of political complexity (Currie et al., 2010), investigated the descent of Acheulian handaxes (Lycett, 2009), and probed behavioural variation among groups of wild chimpanzees (Lycett et al.,

2009). These are just some select studies revealing the diversity of applications on offer (cf. Gray et al., 2007).

Dunn et al. (2011) take data from idiographic studies of languages including typological data and data on how languages are related and combine them using phylogenetic techniques which allow the data to be compared cross-culturally in a controlled way. They tested the nomothetic claim by Greenberg (1963) that languages tended to respect certain word-order constraints, based on the frequencies of certain word-order combinations (e.g. the vast majority of languages with SOV word order are postpositional). This was has been taken as evidence for language universals (e.g. Chomsky, 1986). However, controlling for the effects of descent, Dunn et al. show that the variation in word order universals is better explained by their individual lineages rather than universals over whole language families. For instance, the strengths of different constraints were markedly different between language families. That is, language structure may be influenced more by cultural factors – and the way societies interact – than universal 'innate' biases. This kind of control was not possible before phylogenetic techniques.

Phylogenetic techniques can also be used to address the relationship between language and society more directly: Gray et al. (2009) conduct a phylogenetic analysis of cognates in Austronesian languages to reveal the likely tree of descent of languages and the times at which they split. The results support a 'pulse-pause' model of human migration (Diamond and Bellwood, 2003; Welsch et al., 1992) – the dates at which the linguistic divergence accelerates coincides with innovations in sailing technology which allowed populations to reach more distant islands. Lee and Hasegawa (2011) also use phylogenetic methods to date the origins of Japonic languages and dialects and find evidence that supports an agricultural expansion scenario for populations in Japan.

Phylogenetic techniques are becoming more commonplace in linguistic studies, especially those with large cross-linguistic corpora. They provide a tool for the rigorous testing of specific hypotheses in a controlled way, addressing some of the challenges of the nomothetic approach discussed in this paper.

10. Conclusion

Cross-linguistic data has become more widely available and new statistical approaches provide powerful tools for generating testable hypotheses. However, the data that is most easily accessible and manipulatable is not necessarily the best quality or most relevant. Some formalist approaches to grammar have been criticised for having a weak empirical basis, relying instead on the intuitions of the researchers (e.g. Ferreira, 2005; Wasow and Arnold, 2005; Bresnan, 2007). Similarly, the ease of access and speed of analysis afforded by new data sources and statistical techniques may be in danger of facilitating what Fillmore (1992) called 'computer-aided armchair linguistics'. That is, large-scale statistical analyses can

be used to support an intuition when they should more properly be used to test a specific hypothesis. When conducting a nomothetic study, it's important to use appropriate statistical controls, look carefully at the assumptions being made and recognise the possible weakness of the explanatory power.

The take-home message is that both social structure and language structure are historically contingent and display general patterns. As such, the goal of any research in this field should be testing multiple, and at times contradictory, hypotheses to get a full appreciation of what is being observed in the world. We can probe for lineage-specific (Dunn et al., 2011) and broadscale patterns (Lupyan and Dale, 2010), but there is unlikely to be a one-story-fits-all package at the end of these investigations. The size of phoneme inventories are one such instance: we might find general trends in the data, yet there are likely to be many path-dependent scenarios of development. What we need theoretically is a broad set of underlying mechanisms and how these operate given different demographic conditions. We can investigate this through laboratory experiments and computer simulations. Still, even with support from these methods, the results need to match real-world data: it is here where nomothetic and phylogenetic approaches are most effective. The best way to evaluate the claims of a theory is through independent lines of inquiry that build up a body of mutually supporting evidence. The success of evolutionary theory has been achieved through such degrees of substantiation (Mesoudi et al., 2006). Any approach to language structure and social structure needs to build up a substantial body of data that is informed by a variety of methodological approaches and underpinned by a secure theoretical framework.

Acknowledgements

References

Atkinson, Q. (2011a). Linking spatial patterns of language variation to ancient demography and population migrations. *Linguistic Typology*, 15 (2), 321-332.

Atkinson, Q.D. (2011b). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 332 (6027), 346-349.

Atkinson, Q. (2012). Response to comments on phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 335 (6069), 657-657.

Atkinson, Q.D., Meade, A., Venditti, C., Greenhill, S.J., & Pagel, M. (2008). Languages evolve in punctuational bursts. *Science*, 319 (5863), 588.

Backhaus, J. & Junghanns, K. (2006). Daytime naps improve procedural motor memory. *Sleep Medicine*, 7, 508-512.

Beckner, C., Blythe, R., Bybee, J., Christiansen, M.H., Croft, W., Ellis, N.C., Holland, J., Ke, J., Larsen-Freeman, D., & Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59, 1-26.

Bialystok, E. & Miller, B. (1999). The problems of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition*, 2 (2), 127-145.

Bickel, B. & Nichols, J. (2008). Inflectional synthesis of the verb. In M.S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online* (chapter 22). Munich: Max Planck Digital Library.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44 (2), 235-249.

Bloomfield, L. (1933). *Language.* New York: Henry Holt.

Bogunã, M., Pastor-Satorras, R., Díaz-Guilera, A., & Arenas, A. (2004). Models of social networks based on social distance attachment. *Physical Review E*, 70 (5), 056122.

Bornstein, M. (1973). Color vision and color naming: A psychophysiological hypothesis of cultural difference. *Psychological Bulletin*, 80 (4), 257-285.

Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In S. Featherston & W. Sternefeld (Eds.), *Roots: Linguistics in search of its evidential base* (pp. 75-96). Berlin: Mouton de Gruyter.

Burkett, D. & Griffiths, T. (2010). Iterated learning of multiple languages from multiple teachers. In A. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *Evolution of language: Proceedings of the 8th International Conference EVO-LANG 8* (pp. 58-65). Singapore: World Scientific.

Bybee, J. (2011). How plausible is the hypothesis that population size and dispersal are related to phoneme inventory size? introducing and commenting on a debate. *Linguistic Typology*, 15 (2), 147-153.

Carstairs-McCarthy, A. (2005). The evolutionary origin of morphology. In M. Tallerman (Ed.), *Language origins: Perspectives on evolution* (pp. 166-184). Oxford: Oxford University Press.

Central Intelligence Agency (2009). Sex Ratio. In *The World Factbook 2009.* Washington, DC: Central Intelligence Agency.

Chambers, J.K. (1995). *Sociolinguistic theory: Linguistic variation and its social significance.* Cambridge, MA: Blackwell.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Christiansen, M.H. & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31 (5), 489-508; discussion 509-58.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and non-native language processing. *Language Learning*, 60 (1), 21-43.

Croft, W. (2000). *Explaining language change: an evolutionary approach.* Harlow, Essex: Longman.

Croft, W. (2010). The origins of grammaticalization in the verbalization of experi-
   ence. *Linguistics*, 48 (1), 1-48.

Crop Protection Compendium (2008). Acacia confusa. online, accessed 18-04-2011.
   http://www.cabi.org

Currie, T.E., Greenhill, S.J., Gray, R.D., Hasegawa, T., & Mace, R. (2010). Rise and
   fall of political complexity in island south-east Asia and the Pacific. *Nature*,
   467 (7317), 801-804.

Cysouw, M., Dediu, D., & Moran, S. (2012). Comment on "Phonemic Diversity
   Supports a Serial Founder Effect Model of Language Expansion from Africa".
   Science, 335 (6069), 657.

Dahl, O. (2004). *The growth and maintenance of linguistic complexity*. Amsterdam:
   Benjamins.

de Boer, B. (2000). Self-organization in vowel systems. *Journal of Phonetics*, 28 (4),
   441-465.

DeGraff, M. (2001). On the origin of creoles: A cartesian critique of "neo" darwinian
   linguistics. *Linguistic Typology*, 5, 213-311.

Diamond, J. & Bellwood, P. (2003). Farmers and their languages: The first expan-
   sions. *Science*, 300 (5619), 597-603.

Digiulio, D.V., Seidenberg, M., Oleary, D.S., & Raz, N. (1994). Procedural and de-
   clarative memory: A developmental study. *Brain and Cognition*, 25 (1), 79-91.

Dryer, M.S. (2011). Order of subject, object and verb. In M.S. Dryer & M. Haspelmath
   (Eds.), *The World Atlas of Language Structures Online* (chapter 81). Munich:
   Max Planck Digital Library.

Dunbar, R. (1993). Coevolution of neocortical size, group-size and language in
   humans. *Behavioral and Brain Sciences*, 16 (4), 681-694.

Dunn, M., Greenhill, S.J., Levinson, S.C., & Gray, R.D. (2011). Evolved structure of
   language shows lineage-specific trends in word-order universals. *Nature*, 473
   (7345), 79-82.

Eubank, L. & Gregg, K. (1999). Critical periods and (second) language acquisition:
   Divide et impera. In D. Birdsong (Ed.), *Second language acquisition and the
   critical period hypothesis* (pp. 65-99). Mahwah, NJ: Erlbaum.

Fagyal, Z., Escobar, A., Swarup, S., Gasser, L., & Lakkaraju, K. (2010). Centers and
   peripheries: Network roles in language change. *Lingua*, 120 (8), 2061-2079.

Ferdinand, V. & Zuidema, W. (2009). Thomas' Theorem meets Bayes' Rule: a model
   of the iterated learning of language. In N. Taatgen & H. van Rijn (Eds.), *Proceed-
   ings of the 31st Annual Conference of the Cognitive Science Society CogSci'09*
   (pp. 1786-1791). Wheat Ridge, CO: Cognitive Science Society.

Ferreira, F. (2005). Psycholinguistics, formal grammars, and cognitive science. *The
   Linguistic Review*, 22, 365-380.

Fillmore, C.J. (1992). 'Corpuslinguistics' or 'computer-aided armchair linguistics'.
   In J. Svartvik (Ed.), *Directions in corpus linguistics. Proceedings of Nobel Sym-
   posium 82, Stockholm* (pp. 35-60). Berlin: Mouton de Gruyter.

Garrod, S. (2010). Can iterated learning explain the emergence of graphical symbols? *Interaction Studies*, 11, 33-50.

Gordon, R. (2005). *Ethnologue: Languages of the world* (15th edition). Dallas: SIL International.

Gray, R.D., Drummond, A.J., & Greenhill, S.J. (2009). Language phylogenies reveal expansion pulses and pauses in pacific settlement. *Science*, 323 (5913), 479-483.

Gray, R.D., Greenhill, S.J., & Ross, R.M. (2007). The pleasures and perils of darwinizing culture (with phylogenies). *Biological Theory*, 2 (4), 360-375.

Greenberg, J.H. (1963). *Universals of language: report of a conference held at Dobbs Ferry, New York, April 13-15, 1961.* Cambridge, Mass.: M.I.T. Press.

Greenhill, S.J., Currie, T.E., & Gray, R.D. (2009). Does horizontal transmission invalidate cultural phylogenies? *Proceedings of the Royal Society B: Biological Sciences*, 276 (1665), 2299-2306.

Griffiths, T., Christian, B., & Kalish, M. (2010). Using category structures to test iterated learning as a method for identifying inductive biases. *Cognitive Science*, 32 (1), 68-107.

Hartshorne, J. & Ullman, M. (2006). Why girls say 'holded' more than boys. *Developmental Science*, 9 (1), 21-32.

Haspelmath, M., Dryer, M.S., Gil, D., & Comrie, B. (2008). World Atlas of Language Structures, volume Available online at http://wals.info/feature/22. Accessed on 2011-04-18. Munich: Max Planck Digital Library.

Hawks, J. (2008). From genes to numbers: effective population sizes in human evolution. In J.-P. Bocquet-Appel (Ed.), *Recent advances in palaeodemography: Data, techniques, patterns* (pp. 9-30). Dordrecht: Springer.

Hay, J. & Bauer, L. (2007). Phoneme inventory size and population size. *Language*, 2, 388-400.

Henrich, J. (2004). Demography and cultural evolution: How adaptive cultural processes can produce maladaptive losses: The tasmanian case. *American Antiquity*, 69 (2), 197-214.

Hunley, K., Bowern, C., Healy, M., Hunley, K., Bowern, C., & Healy, M. (2012). Rejection of a serial founder effects model of genetic and linguistic coevolution. *Proceedings of the Royal Society B: Biological Sciences*, 279 (1736), 2281-2288.

Hurford, J.R. (1991). The evolution of the critical period for language acquisition. *Cognition*, 40 (3), 159-201.

Hyltenstam, K. & Abrahamsson, N. (2003). Maturational constraints in SLA. In C.J. Doughty & M.H. Long (Eds.), *The handbook of second language acquisition* (pp. 539–588). Malden, MA: Blackwell.

International Monetary Fund. Online (2011). World economic outlook, April 2011.

Isbell, L.A. & Young, T. (1996). The evolution of bipedalism in hominids and reduced group size in chimpanzees: Alternative responses to decreasing resource availability. *Journal of Human Evolution*, 30, 389-397(9).

Jaeger, T., Graff, P., Croft, W., & Pontillo, D. (2011). Mixed effect models for genetic

and areal dependencies in linguistic typology. *Linguistic Typology*, 15, 281.

Johnson, J. & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, 21 (1), 60-99.

Kaestle, F.A. & Smith, D.G. (2001). Ancient mitochondrial DNA evidence for prehistoric population movement: the numic expansion. *American Journal of Physical Anthropology*, 115 (1), 1-12.

Kay, P. & Maffi, L. (2008). Number of basic colour categories. In M. Haspelmath, M.S. Dryer, D. Gil, & B. Comrie (Eds.), *The World Atlas of Language Structures Online* (chapter 133). Munich: Max Planck Digital Library.

Kay, P. & Regier, T. (2003). Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences*, 100, 9085-9089.

Kellerman, E. & Sharwood Smith, M. (Eds.) (1986). *Crosslinguistic influence in second language acquisition*. Oxford: Pergamon Press.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (31), 10681-10686.

Kirby, S., Dowman, M., & Griffiths, T.L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of the United States of America*, 104 (12), 5241-5245.

Kirby, S. & Hurford, J. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In A. Cangelosi & D. Parisi (Eds.), *Simulating the evolution of language* (pp. 121-148). London: Springer.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277 (5326), 684-686.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Lee, S. & Hasegawa, T. (2011). Bayesian phylogenetic analysis supports an agricultural origin of Japonic languages. *Proceedings of the Royal Society B: Biological Sciences*, 278 (1725), 3662-3669.

Levinson, S. (2006). Introduction: The evolution of culture in a microcosm. In S.C. Levinson & P. Jaisson (Eds.), *Evolution and culture: A Fyssen Foundation Symposium* (pp. 1-41). Cambridge: MIT Press.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12 (3), 251-285.

Lupyan, G. & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, 5 (1), e8559.

Lycett, S. & Norton, C. (2010). A demographic model for palaeolithic technological

evolution: the case of east asia and the movius line. *Quaternary International*, 211 (1), 55-65.

Lycett, S.J. (2009). Understanding ancient hominin dispersals using artefactual data: a phylogeographic analysis of acheulean handaxes. *PLoS One*, 4 (10), e7404.

Lycett, S.J., Collard, M., & McGrew, W.C. (2009). Cladistic analyses of behavioural variation in wild pan troglodytes: exploring the chimpanzee culture hypothesis. *Journal of Human Evolution*, 57 (4), 337-349.

MacKeigan, T. & Muth, S.Q. (2006). A grammatical network of Tzotzil Mayan colour terms. In C.P. Biggam & C.J. Kay (Eds.), *Progress in Colour Studies. Volume I: Language and culture* (pp. 23-36). Amsterdam: John Benjamins.

Maddieson, I. (2008). Tone. In M.S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online* (chapter 13). Munich: Max Planck Digital Library.

Matisoff, J. (1978). *Variational semantics in Tibeto-Burman: The 'organic' approach to linguistic comparison.* Philadelphia: Institute for the Study of Human Issues.

McPherson, J.M., Smith-Lovin, L., & Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415-444.

McWhorter, J. (2008). Why Does a Language Undress? Strange Cases in Indonesia. In M. Miestamo, K. Sinnemaki, & F. Karlsson (Eds.), *Language complexity: Typology, contact, change* (pp. 167-190). Amsterdam: John Benjamins.

Mesoudi, A. & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B*, 363, 3489-3501.

Mesoudi, A., Whiten, A., & Laland, K.N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, 29 (4), 329-47; discussion 347-383.

Milroy, J. & Milroy, L. (1985). *Authority in language.* London: Routledge and Kegan Paul.

Nerbonne, J. (2010). Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365 (1559), 3821-3828.

Nettle, D. (1999). *Linguistic diversity.* Oxford: Oxford University Press.

Nettle, D. & Dunbar, R.I.M. (1997). Social markers and the evolution of reciprocal exchange. *Current Anthropology*, 38 (1), 93-99.

Newman, M.E.J. & Park, J. (2003). Why social networks are different from other types of networks. *Physical Review E*, 68 (3), 036122.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism.* Amsterdam: John Benjamins..

Pericliev, V. (2004). There is no correlation between the size of a community speaking a language and the size of the phonological inventory of that language. *Linguistic Typology*, 8, 376-383.

Quillinan, J., Kirby, S., & Smith., K. (2010). Co-evolution of language and social network structure through cultural transmission. In A.D. Smith, M. Shouwstra, B. de Boer, & K. Smith (Eds.), *Proceedings of the 8th International Conference on the Evolution of Language EVOLANG 8* (pp. 475-476). Singapore: World Scientific.

Reali, F. & Griffiths, T. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, 111 (3), 317-328.

Richerson, P.J., Boyd, R., & Bettinger, R.L. (2009). Cultural innovations and demographic change. *Human Biology*, 81 (2-3), 211-235.

Roberts, G. (2010). An experimental study of social selection and frequency of interaction in linguistic diversity. *Interaction Studies*, 11 (1), 138-159.

Roberts, S. & Winters, J. (2012). Constructing knowledge: Nomothetic approaches to language evolution. In L. McCrohon, T. Fujimura, K. Fujita, R. Martin, K. Okanoya, R. Suzuki, & N. Yusa (Eds.), *Five approaches to language evolution: Proceedings of the Workshops of the 9th International Conference on the Evolution of Language* (pp. 148-157). Singapore: World Scientific.

Sapir, E. (1912). Language and environment. *American Anthropologist*, 14 (2), 226-242.

Schachter, J. (1974). An error in error analysis. *Language Learning*, 24 (2), 205-214.

Scott-Phillips, T.C. & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, 14 (9), 411-417.

Séguy, J. (1973). La dialectom´etrie dans l'atlas linguistique de gascogne. *Revue de Linguistique Romane*, 37 (145), 1-24.

Shennan, S. (2000). Population, culture history and the dynamics of culture change. *Current Anthropology*, 41, 811-835.

Smith, C. & Smith, D. (2003). Ingestion of ethanol just prior to sleep onset impairs memory for procedural but not declarative tasks. *Sleep*, 15 (26), 185-191.

Smith, K. & Kirby, S. (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363 (1509), 3591-3603.

Smith, K. & Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. *Cognition*, 116 (3), 444-449.

Tamariz, M., Cornish, H., Roberts, S., & Kirby, S. (2012). The effect of generation turnover and intelocutor negotiation on linguistic structure. In T.C. Phillips, M. Tamariz, E.A. Cartmill, & J.R. Hurford (Eds.), *The evolution of language: Proceedings of the 9th International Conference EVOLANG9* (pp. 555-556). Singapore: World Scientific.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, P. (2004). Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology*, 8, 305-320.

Trudgill, P. (2011). Social structure and phoneme inventories. *Linguistic Typology*, 15, 155-160.

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141-178). Washington, DC: Georgetown University Press.

Vogt, P. (2009). Modelling interactions between language evolution and demography. *Human Biology*, 81 (2-3), 237-258.

Wasow, T. & Arnold, J. (2005). Intuitions in linguistic argumentation. *Lingua*, 115, 1481-1496.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. W.P. Lehmann, Y. Malkiel (Eds.), *Directions for historical linguistics: A symposium* (pp. 97-195). Austin: University of Texas Press.

Welsch, R.L., Terrell, J., & Nadolski, J.A. (1992). Language and culture on the north coast of new guinea. *American Anthropologist*, 94 (3), 568-600.

Wichmann, S. & Holman, E.W. (2009). Population size and rates of language change. *Human Biology*, 81 (2-3), 259-74.

Wichmann, S., Rama, T., & Holman, E. (2011). Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguistic Typology*, 15, 177-197.

Wikipedia (2011). List of flags by number of colors. online, accessed 12-04-2011. http://en.wikipedia.org/wiki/list of flags by number of colors.

Winters, J. (2010). Phoneme inventory size and demography. In Proceedings of the 24th Language at Edinburgh Lunch, Edinburgh.

World Values Survey Association (2009). World Values Survey 1981-2008 Official Aggregate v. 20090901. Madrid: ASEP/JDS.

Wray, A. & Grace, G.W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, 117, 543-578.