that these relationships are often coded overtly in the verbal morphology of languages. In addition to the languages discussed in this chapter, this system has also been shown to be central to the organization of the verbal systems of the following languages: Tagalog (Foley and Van Valin, 1984), Bonggi (Boutin, 1994), Yatye (Kwa, Nigeria; Foley and Van Valin, 1984; Stahlke, 1970), Italian (Centineo, 1996), Icelandic (Van Valin, 1991), Croatian (Dahm-Draksic, 1998), Korean (B. Yang, 1994), Japanese (Hasegawa, 1996; Toratani, 1998), and Bribri (Chibchan, Costa Rica, Tomcsányi, 1988). Weist (2002) shows that these distinctions are fundamental to the acquisition of verbal systems. Thus, the *Aktionsart* distinctions discussed in this chapter are prime candidates for semantic universals.

8   Language change and universals

Joan Bybee

1   Introduction

As traditionally understood, universals of language are cross-linguistic generalizations concerning synchronic grammars, and their explanations usually appeal to functional principles thought of in a synchronic domain. It stands to reason, however, that any synchronic pattern must have a diachronic dimension, since that pattern had to come into being in some way. One could even argue, as I did in Bybee (1988), that we cannot be sure of the validity of a functional explanation for a synchronic universal unless we can confirm that that functional consideration was applicable in the formation of the synchronic pattern. That is, all explanations of synchronic universals must have a diachronic dimension.

In the current chapter, I outline a position on the role of diachrony in universals, whose logical consequence is that the true universals of language are not synchronic patterns at all, but the mechanisms of change that create these patterns. This position is an extension of the theory of diachronic typology formulated and practiced by Joseph H. Greenberg, to whom this chapter is dedicated.[1]

In several papers, Greenberg proposed a method for the study of typology and universals which he called dynamic comparison or diachronic typology.[2] In this method, typological patterns are shown to emerge from common diachronic changes that arise in related and unrelated languages. It has become clear subsequently that what Greenberg elaborated in the many domains of language that he studied was not a comparative methodology so much as a theory of language that has great potential for explanation. His so-called method is in fact a model of linguistic evolution and change in which the grammars of individual languages are emergent from the processes of change that are operative in all languages at all times. In this view, the true universals of language are

---

[1] Other work formulated in this framework is found, for example, in Givón (1979) and Croft *et al.* (1990).

[2] This method/theory is evident in much of Greenberg's work, but Greenberg (1969) contains a particularly explicit statement of the theory, and Greenberg (1978b) illustrates the method for grammatical comparison.

the mechanisms of change that propel the constant creation and re-creation of grammar.

In this chapter, I will begin with a brief review of the theoretical status of universals in modern theories, and then briefly demonstrate Greenberg's method of dynamic comparison. I will then apply this theory to two examples: the grammaticization of tense and aspect in the languages of the world and the distribution of the phoneme /h/ cross-linguistically. In both cases, I will present the common paths of change for these elements and argue that the diachronic paths present much stronger cross-linguistic patterns than any comparison based solely on synchronic grammars. However, the story does not end here. Behind these paths of change are common mechanisms of change that occur in the process of language use and these mechanisms must be carefully examined, as they are the true universals of language.

2   Observations of cross-linguistic regularities

There are several levels at which universals of language can be formulated and incorporated into a linguistic theory. The first and most basic level is the level of observation. By comparing a number of distinct languages we can come up with a list of properties that all languages share, such as (1)–(5):

Phonology:
(1) All languages have vowels and consonants.
(2) All languages have stop consonants.
(3) All languages have a low vowel, [a]

Morphosyntax:
(4) All languages have nouns and verbs.
(5) All languages have a negative construction.

A major problem we confront immediately is that absolute universals are disappointingly few. Missing from the lists above are any statements about the morphological categories that languages have. We cannot truthfully say that all languages have tense, aspect, or mood expressed morphologically (Bybee, 1985).

Another approach is to formulate relative statements, using the notion of *markedness*. As this term has come to be used today, it distinguishes the more usual or common member of an opposition, the unmarked member, from the less usual, less common member, the marked member. Thus a statement such as (6) can be formulated:

(6)   Oral vowels are unmarked compared to nasal vowels.

This statement covers the following facts (Greenberg, 1966):

(7)   Unmarked sounds are more common as phonemes in the languages of the world.

(8)   If a language has the marked sound as a phoneme, it also has the unmarked one as a phoneme.

(9)   Unmarked sounds occur more frequently in running text than the marked ones, in languages that have both.

All of the statements given so far are observations about what occurs in the languages of the world. No explanatory account of how or why these regularities manifest themselves in individual languages is provided by these statements. A fully elaborated theory of language must provide such an account.

3   Generative theories

Within generative theory, it is proposed that the observational statements of cross-linguistic regularities are built directly into the synchronic grammar of each language. As universals, they are taken to be part of the innate apparatus that children bring to the task of language acquisition. The Marking Conventions of Chomsky and Halle (1968) supply the unmarked value of features. Thus, the unmarked voiceless obstruents and oral vowels have their features [–voice] and [–nasal] supplied by the Marking Conventions, implying that children do not have to learn these feature values and that these values do not add complexity to the grammar. They are "free" because they are provided by universals.

Similarly, Optimality Theory (OT), which developed out of generative theory, formulates constraints as universals and hypothesizes that such constraints are given innately, and only their language-particular ordering needs to be learned by children (Prince and Smolensky, 1997). For example, a constraint formulated to account for (6) would say simply:

(10)   Vowels are [–nasal]

The facts in (6–9) are distilled into this one statement, which will be allowed to apply in derivations in languages with only oral vowels, but it will be overridden in languages with nasal vowels.

What generative theory and OT have in common is that the empirical observations about properties of synchronic states in the languages of the world are directly inserted (sometimes in a simplified form) into the innate apparatus of Universal Grammar, without any attempt being made to formulate them in such a way as to explain their existence. To say that a constraint is universal and innate is to isolate it from possible explanation on the basis of factors outside of language, and, indeed, to preclude the need for further explanation. Even those

versions of OT that appeal to phonetic or functional explanations for constraints fail to provide the crucial diachronic link between the functional factor and the existing synchronic state.

In my opinion, such theories are rather primitive and they entirely miss the chance to provide subtle and sensitive explanations for the way the cross-linguistic regularities cited above arise. A much more elegant theory, and one that accounts for the fact that there are so few absolute universals, is that developed in the work of Joseph Greenberg, to which we now turn.

4   Diachronic typology

In Greenberg's approach, language states are the product of change. Cross-linguistic comparison is informed by the idea that synchronic states represent different stages of development of subsystems of grammar. For instance, the fact that all languages have oral vowel phonemes and only some languages have nasal vowel phonemes is due to the fact that nasal vowels develop out of oral vowels in the context of a nasal consonant, which subsequently is lost. In a final stage, vowel nasality can also be lost, taking us back to the original situation – a language with only oral vowels. The most basic path of change is (Greenberg, 1978a):

(11) VN $>$ $\tilde{V}N$ $>$ $\tilde{V}$ $>$ V
        I   II   III   IV

A language in stages I, II, or IV has only oral vowel phonemes. Only in stage III are there nasal vowel phonemes. Since at all stages oral vowels exist, the statements in (7) and (8) are true: all languages have oral vowel phonemes and languages with nasal vowels also have oral vowels. In general then, nasal vowels are less common in the languages of the world because there is only one way for them to develop: from oral vowels in a restricted context. For the same reason, statement (9), which other theories cannot explain, is true: nasal vowels will be less frequent in running discourse than oral vowels. A theory that simply states that nasal vowels are marked misses the subtlety of this situation and the diachronic explanation for it.

In "How does a language acquire gender markers?" Greenberg (1978b) postulates a universal diachronic sequence, based on numerous examples from related and unrelated languages, that leads from demonstrative to definite article to non-generic article to Noun marker (a classificatory marker or gender).

(12)   DEMONSTRATIVE > DEFINITE ARTICLE > NON-GENERIC
        ARTICLE > NOUN MARKER

Without the discovery of this diachronic trajectory or path, languages with each of these Noun treatments might be considered as constituting independent types with no basis for comparison. Greenberg's brilliant discovery was that, in unrelated families and languages, parts or all of this path of change are attested, and, furthermore, that the progression along the path is unidirectional. The changes always move in the direction of the arrows, and never in reverse.

The relation among these structures, when only viewed synchronically, is comparable to a purely synchronic view of an acorn, an oak seedling, a full grown oak tree, and wood products, such as lumber. As individual entities they appear to have little in common, but on a developmental account, one becomes the other in a unidirectional fashion. Linguistic theories that just compare synchronic states are like a hypothetical biological theory that compared the seeds of various plants with no regard to what type of developmental trajectory a seed is meant to embark upon. Such a theory would also study a seedling without considering its past and future development. Such a theory of biology would have little chance of explaining the properties of seeds or seedlings or full-grown trees and the wood products that result from them. Similarly, a linguistic theory that compares only the synchronic states of languages is unlikely to hit upon valid explanations for the diversity of states attested in the languages of the world.

5   Grammaticization

The stages on the path in (12) are related to one another by the process of grammaticization. Grammaticization is the process by which constructions arise in languages and the lexical items in them become grammatical morphemes (Bybee *et al.*, 1994; Heine *et al.*, 1991; Hopper and Traugott, 1993; Meillet, 1958 [1912]). The diachronic path in (12) starts with grammatical morphemes (demonstratives) that then become more grammatical. But there are many cases in which the lexical source of a grammatical morpheme can be identified. Recent research has identified numerous paths of grammaticization that are known to have occurred in many unrelated languages. Bybee *et al.* (1994), and Heine and Kuteva (2002) document the cross-linguistic validity of numerous paths of grammaticization.

For instance, in Bybee *et al.* (1994), we used a cross-linguistic sample of seventy-six languages, selected to be maximally unrelated genetically, to study the diachronic origins of tense, aspect, and mood markers. The three paths of change, I–III, for tense and aspect, first proposed in Bybee and Dahl (1989), were robustly evidenced in these unrelated languages, although other, less common

paths, are also found. The categories named on these paths, such as anterior, perfective, etc., have also been verified as cross-linguistically applicable in Bybee (1985) and Dahl (1985).

I. The perfective path

(i) "be," "have" + PP > RESULTATIVE
(ii) "come (from)" } > ANTERIOR > PERFECTIVE/PAST
(iii) "finish" > COMPLETIVE

Examples
(i) A stative auxiliary plus a past participle for anterior and past/perfective is documented in Romance and Germanic.

 (13) French
  il a chanté.
  he has sing–PP
  "He sang."

(ii) "Come from" develops into anterior and past in dialects of Atchin (Oceanic) (Bybee *et al.*, 1994), Teso, and Jiddu (Heine and Reh, 1984).

 (14) Jiddu (Heine and Kuteva, 2002, 72)
  y- aam-ooku
  3:M eat-come
  "He has just eaten."

(iii) Many languages have anteriors from "finish" (e.g. Bongu, Temne, and Lao, see Bybee *et al.*, 1994), but the entire path can be reconstructed in Bantu languages (Voeltz, 1980).

II. The present/imperfective path

(i) "be located at" }
(ii) "movement while" } > PROGRESSIVE > PRESENT/IMPERFECTIVE
(iii) reduplication }

Examples
(i) Heine *et al.* (1991) report over 100 African languages with a locative source for progressive.

 (15) Godié (Marchese, 1986, 63)
  ɔ kù ɓlɪ-dʌ.
  he be. at sing-place
  "He is singing."

(ii) Movement sources for progressives are found in Spanish, Tojolabal, Tok Pisin, and others (see Bybee *et al.*, 1994).

 (16) Spanish
  Andaba escribiendo para los periódicos.
  go:IMPF:3S write:GER for the newspapers
  "He was writing for the newspapers."

(iii) Reduplication starts out as iterative, then can become progressive and eventually imperfective, e.g. in Nakanai, Rukai, Gugu-Yalanji, and Trukese (Bybee *et al.*, 1994).
  Nakanai (Johnston, 1980)

 (17) *Eia o-io sa-sapa.*
  3S at-there REDUP-sweep
  "She is there sweeping."

 (18) *Eia sa-sapa te la kavikoki.*
  3S REDUP-sweep PREP NM morning
  "She sweeps in the mornings."

(iv) Progressives have become imperfectives in Yoruba, Scots Gaelic, and Turkish (Comrie, 1976); the progressive has become a present in the Chamus dialect of Maa.

III. Future

(i) "want" }
(ii) "movement towards" } > INTENTION > FUTURE
(iii) "soon," "after" }

Examples
(i) Desire futures occur in English, Danish, Dakota, Serbo-Croatian, Swahili, etc. (see Bybee and Pagliuca, 1987).

 (19) I'll try to find it.   *'ll < will < willan* "to want"

(ii) Movement futures occur in English, Spanish, Zulu, and many other African languages (Heine and Kuteva, 2002).

 (20) *Va a llover.*
  "It's going to rain."

(iii) Temporal Adverbs give rise to futures in Trukese, Bari, Chepang, and Tok Pisin (Bybee *et al.*, 1994).

(21) Tok Pisin: *by and by* > *baimbai* > *bai* (Romaine 1995)
    em bai tupela sindaun.
    "The two of them will sit down."

(iv) An intermediate stage of *intention*, especially of first person subjects, is
    well documented for all future sources (Bybee *et al.*, 1994).

This remarkable similarity in grammaticization paths across unrelated languages strongly suggests that universals of diachronic development be included in a theory of language universals. I would further argue that the diachronic universals are much stronger than any synchronic universals we can formulate concerning the presence and meaning of tense and aspect markers in the languages of the world. Given the seven common category types involved in the paths above – anterior, perfective, past, progressive, present, imperfective, and future – a very large number of possible synchronic tense/aspect systems can be derived. The existence of anterior, progressive, and future are independent of any other categories. All or any one of them can occur in a language with or without perfective/imperfective or present/past. Thus it appears that the paths are independent in the sense that a development along one path does not affect a development along other paths. Nor does a development along one path preclude other developments along the same path during the same time period. In fact, there are only two generalizations that we have observed that affect multiple paths in a language.

First, since present, past, perfective, and imperfective tend to be inflectional – meaning that they belong to obligatory categories – if a language has a past tense, then it also has a present tense (sometimes expressed by zero). Similarly, if a language has an imperfective, that implies the existence of a perfective (again, possibly expressed by zero). Languages may also have both types of opposition. In that case, the imperfective has a tense distinction and typically the perfective has past as a default interpretation (Dahl, 1985). However, there do not appear to be any restrictions on which other tense or aspect categories can appear in a language with present/past or perfective/imperfective.

The second restriction has to do with how far grammaticization progresses in languages of different types. As Dahl (1985) discovered, past and perfective tend to have inflectional expression, so an analytic language which lacks inflection is also extremely likely to lack a past or perfective. In Bybee *et al.* (1994) this finding was replicated on a different sample. Bybee (1997) interprets this constraint as owing to the nature of grammaticization in analytic languages. It appears that such languages do not carry grammaticization to the end points in the paths but rather go no farther than, e.g., progressive and anterior, which are common categories in analytic languages. Of course, if all languages carried grammaticization to the fullest extremes, there would be no analytic languages, i.e. no language lacking inflection. As Sapir (1921) pointed out, morphological

typology corresponds to semantic typology. Some languages do not express the very abstract relational meanings (Sapir's "pure relational concepts") with grammatical markers. Bybee (1997) has argued that such meanings, to the extent that they occur at all, arise only by inference in analytic languages.

If these are the only restrictions on grammaticization, then the possible combinations of aspect and tense categories is rather free – a language could have a construction at any point along any of these paths, or it could have multiple constructions traversing the same path, as many languages do. As a result, the possible synchronic combinations of tense/aspect constructions are relatively unconstrained (see Dahl, 1985). Thus the very robust and very specific paths of development shown above constitute much stronger cross-linguistic statements than any statements we could devise about synchronic states.

6   Similarities among the paths of change

Even the earliest work on grammaticization recognizes that the same set of processes is at work in creating these different paths of change. In recent years, researchers have commented on the semantic and pragmatic changes that occur in constructions as they grammaticize. Grammatical meaning is more abstract, more generalized, more subjective and discourse-oriented than lexical meaning. What mechanisms are responsible for the creation of grammatical meaning?

Although there is still much to learn, substantial progress has been made already in the identification and study of some important mechanisms of change. Here I will emphasize the role of repetition in creating the semantic and phonological changes in grammaticization, and the role of pragmatic inference. Other mechanisms of course exist, but my point is to illustrate the universality and basicness of the mechanisms underlying the paths laid out above and not to provide a full account of grammaticization.

6.1   The role of repetition

Grammaticization is always accompanied by a sharp increase in frequency of use of the construction undergoing change. The frequency increase is both a result of the process and a contributor to it, as repetition has certain effects on neuromotor and cognitive representations.

Some recent studies of grammaticization have emphasized the point that grammaticization is the process of automatization of frequently occurring sequences of linguistic elements (Boyland, 1996; Haiman, 1994; Bybee, 2002a). Boyland (1996) points out that the changes in form that occur in the grammaticization process closely resemble changes that occur as non-linguistic skills are practiced and become automatized. With repetition, sequences of units that

were previously independent come to be processed as a single unit or chunk. This repackaging has two consequences: the identity of the component units is gradually lost, and the whole chunk begins to reduce in form. These basic principles of automatization apply to all kinds of motor activities: playing a musical instrument, tying your shoe laces, driving a car. They also apply to grammaticization. A phrase such as *(I'm) going to (V E R B)*, which has been frequently used over the last couple of centuries, has been repackaged as a single processing unit. When a sequence of actions is automated, it gains in fluency. The component gestures (in this case, articulatory) reduce and overlap. Thus *going to* undergoes both vowel and consonant reduction, yielding variants such as [gə̃nə].

Repetition of grammaticizing constructions leads to habituation. The earliest discussion of grammaticization recognized that grammatical morphemes lose components of their original lexical meaning and become much more general and abstract. For instance, *will* loses the volitional aspect of its meaning and *be going to* loses the spatial movement components. This process has been called "bleaching" or "generalization of meaning." The latter term is especially appropriate, because the loss of specificities of meaning makes a morpheme applicable in a wider range of contexts. For example, if *will* does not signal volition, it can be used with a wider range of subjects, including inanimate objects.

Repetition itself diminishes the force of a word, phrase, or construction. Examples are well known: *iterate* doesn't seem to mean "repeat" quite strongly enough, so English-speakers tend to add *re-*; with repetition the strength of that fades and we have to say *reiterate again*. In grammaticization, the generalization or bleaching of the meaning of a construction is caused by frequency, but it also contributes to additional frequency, since a generalized construction can be used in more contexts, and thus the change is propelled along the path, gaining momentum as it goes.

Automatization and habituation through repetition are ongoing in all languages at all times, i.e. they are universally present when people use language. These two processes explain a number of properties of the grammar of human languages: (i) grammatical morphemes are reduced phonologically, and dependent, via affixation and cliticization, on lexical material; (ii) grammatical meaning is highly general and abstract in nature.

6.2   *Pragmatic inferences*

Pragmatic inference is an important mechanism of change in grammaticization. It has been studied intensely by Elizabeth Traugott in various works (e.g. Traugott, 1989; Traugott and Dasher, 2002). It is widely accepted that an important feature of the communication process is the ability to make inferences: the

speaker must be able to judge which details the hearer can supply and formulate his/her utterances accordingly, and the hearer must fill in details not supplied by the speaker. Thus, the hearer is constantly asking "why is s/he asking me or telling me this?" The workings of pragmatic inference in producing semantic change are nicely demonstrated by an example presented in Hopper and Traugott's (1993) book on grammaticization. In Shakespeare's English, *be going to* had its literal meaning of movement in space towards some goal. However, given the apparent interest of human beings in goals and purposes, even in Shakespeare's English, the information value of *be going to* was less about movement in space and more about purpose. Consider example (22):

(22)   *Duke* Sir Valentine, whither away so fast?
        Val. Please it your grace, there is a messenger
                That stays in to bear my letters to my friends,
                And I am going to deliver them.
                            (1595, Shakespeare, *Two Gentlemen of Verona*, III.i.51, in
                                                        Hopper and Traugott, 1993)

Note that even though the Duke asks about movement ("Where are you going so fast?"), what he really wants to know is Valentine's intention or purpose. Note also that, although Valentine answers in terms of movement, he does not say exactly where he is going; rather he states his intention.

When the same pattern of inference occurs frequently with a particular grammatical construction, those inferences can become part of the meaning of the construction. If *be going to* is frequently used to talk about intentions, it begins to have intention as part of its meaning. The literature on grammaticization is full of such instances (Traugott, 1989; Bybee *et al.*, 1994).

Note that common paths of change, such as the development from movement towards a goal to Future would not be attested across languages unless users of these languages made very similar inferences under similar conditions. That is, the repetition across languages of the change in meaning from "movement towards a goal" to "intention" is evidence that speakers in different cultures tend to infer intentions in the same context; similarly, changes from temporal sequence (as English *since*, originally meaning "after the time that") to causation indicate that language users are prone to infer causation.

7   **Synchronic patterns as emergent**

Lindblom *et al.* (1984) explain how a complex system can emerge from the repetition of many local actions by using the example of how termites build a nest. The nest has a complex architecture, with pillars and arches and yet it is built without a master plan. Each termite repeats a simple sequence of actions

that result in the elaborate nest. The termites each carry a small amount of gluti-
nous sand seasoned with pheromones; the termite drops its load of sand when it
detects a concentration of pheromones. If they begin on a flat surface, at first the
deposits are randomly placed. However, very soon local peaks of pheromones
and sand begin to appear and these attract even more deposits. The arches are
formed when two peaks are very close to one another, and the great quantity of
pheromones makes it very likely that more deposits will unite the pillars in an
arch.

The mechanisms discussed here produce change over a long period of time
by operating in real time in individual usage-events. Change by automatization
and inferencing build up over multiple repetitions as language is being used by
individual speakers. These mechanisms operate on language-specific material
in a predictable way, producing the paths of change discussed above. The paths
of change then produce the synchronic linguistic structures. Because they were
produced by the same mechanisms across languages, they resemble one another.
Because they were produced in different languages with different linguistic
material as input to the process, with some differences in the contexts of use,
the outcomes are similar but not identical. As a result, absolute universals of
language are rare. Moreover, certain minor paths of change are also possible:
since only the mechanisms of change are universal, the way they interact with
one another, with language-particular material, and with the social context may
produce variations on these paths of change.

As an example of a minor path of change, consider a development that can
occur in the perfective path, given above in Section 5. Bybee *et al.* (1994) find
that a possible development of a resultative coming from "have" or "be" plus
a participial verb form is an evidential of inference from results or indirect
evidence. This development is found in the GRAMCATS sample in Udmurt
(Uralic), Inuit, and Tucano (Tucanoan), and it is also known to have occurred
in Turkish, Bulgarian, Macedonian, and Georgian. Thus, while it appears that
the most common path for a resultative is to take on anterior functions and
eventually develop into a past or perfective, in these languages the resultative
develops into a past evidential. While such a development might constitute
an exception to the general path of change, the mechanism that produced the
change is one of the well-attested mechanisms: change by pragmatic inference.
If a speaker makes an assertion citing the results of a past action (*Tom is gone*),
the hearer is entitled to infer that the speaker knows the past action largely
through its results (that Tom is no longer here, rather than that the speaker
witnessed Tom's leaving). If such an inference is commonly made, it becomes
part of the meaning and the construction would come to mean "Tom must have
gone" rather than "Tom is gone." From the resultative evidential function, the
construction could generalize to other evidential functions (Bybee *et al.*, 1994,
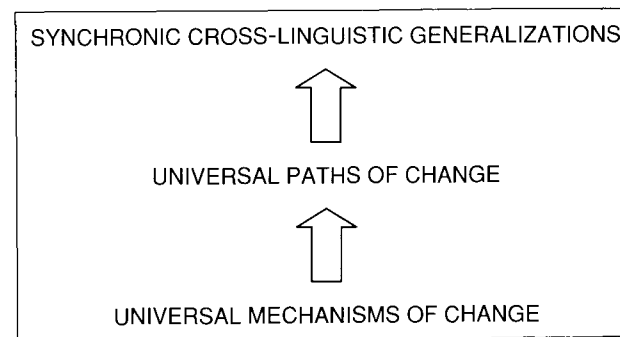95–97). Thus, while this development is somewhat restricted in terms of the

Figure 8.1   The relations among three tiers of cross-linguistic phenomena.

number of languages that undergo it, it is produced by the same mechanisms
as the more widespread changes.

Figure 8.1 presents this view of language universals. At the most basic level
are the mechanisms of change. These mechanisms are operative as language
users produce many local and specific actions in the process of communicating.
The repetition of communicative acts leads to automatization and reduction of
form, habituation and generalization of meaning, as well as the conventional-
ization of pragmatic inference.

These mechanisms create paths of change which are often similar cross-
linguistically. As a by-product of these paths, synchronic states may also bear
some resemblance to one another.

8   Application to phonology: the status of /h/ in the languages of the world

Another area where diachronic universals are stronger than synchronic state-
ments is in phonology. There are general paths of phonological change that
principally lead to weakening and reduction (Browman and Goldstein, 1992;
Mowrey and Pagliuca, 1995). In the case of consonants, it is common for
stops to become fricatives and, further, for fricatives to lose their supraglottal
articulation and end up, in the case of voiceless fricatives, reduced to simple
glottal frication, or [h]. The glottal fricative itself, being weakly articulated and
acoustically less salient, is subject to loss. Such changes are more common in
certain positions in the word and syllable than in others. In general, reduction
is favored in word-final and syllable-final position, which are characterized by
less articulatory force.

It is important to note that the implementation of any sound change is a com-
plex phenomenon, involving differential progress of the change in phonetic and

lexical environments and different effects of the change in terms of contrasting segments. These factors are extremely important for an eventual understanding of how and why sound change takes place, as well as for the eventual outcome of the change. However, if we abstract the general patterns from these factors which will differ for each sound change, universal paths of change are evident since the underlying forces that cause change are the same across languages. The particular changes discussed here are due to the reduction of gestures that occurs when speech becomes more and more automated with use.[3]

Commonly reported changes creating and subsequently eliminating [h] are seen in (23–25).[4]

(23)   p > f > h > Ø

Parts of this path (23) are documented in different languages: Japanese has undergone a change that reduced all prevocalic instances of /p/ to a fricative that assimilates to the place of articulation of the following vowel – [ha], [çi], [ɸɯ], [he], [ho]. The frequently reported absence of /p/ in the languages of the world is very likely due to the propensity of /p/ to undergo weakening (Maddieson, 1984, 36–37), which in turn has a phonetic explanation. Spanish and other Romance languages have undergone a change that reduced word-initial [f] to [h] and further to Ø.

The path in (24) is also common and it converges with (23) (just as grammaticization paths converge) to create [h] which eventually deletes.

(24)   s > h > Ø

Such changes are amply documented in syllable-final position in dialects of Spanish in Andalusia, the Caribbean, and other Latin American dialects. Word-initial [s] also weakens to [h] in Ancient Greek (Méndez Dosuna, 1996) and in Mexican dialects of Spanish.

Finally, the path in (25) is documented, though less commonly as velar fricatives in general are less common in the languages of the world (Maddieson, 1984). This path creates a [h] as do the others, and it is also subject to deletion.

(25)   x > h > Ø

Instances of /h/ in Germanic languages correspond to Proto-Indo-European *k (Latin *cord*-; English *heart*) and are postulated to have evolved through a velar fricative stage. Indeed the velar fricative occurs in Old and Middle English,

---

[3] The evidence that phonetically conditioned sound change is a part of the continuing automation process is the fact that many sound changes show evidence of lexical diffusion from the most frequent words to the least frequent (Phillips, 1984; Bybee, 2002b). On some of the complexity of the changes discussed below, see Pensado (1993), Méndez Dosuna (1996).

[4] These examples, the cross-linguistic survey, and the general theoretical framework are taken from Bybee (2001).

---

as it does in German *Nacht* "night" (cf. Latin *noct*-). In some dialects of British English, initial /h/ is now deleting. In southern Spain and American dialects of Spanish, an earlier /x/ is produced as [h] (Penny, 1991).

Thus, all three converging paths are attested in independent cases, and I suspect that a concentrated study would turn up many more cases in more widely dispersed languages.[5] The cross-linguistic validity of these paths of change, then, is secure. However, any synchronic statements that we might attempt for the status of /h/ in phoneme inventories do not succeed. Synchronically, we find languages in which /h/ is restricted to syllable-initial or word-initial position, and we find languages in which /h/ is favored in syllable-final position. American English is an example of the first type, where /h/ does not occur finally, but does occur word-initially (*house* [haus]), and syllable-initially before a stress (*vehicle* ['viɪkl] vs. *vehicular* [vɪ'hɪkjulr]). Certain dialects of Spanish provide examples of the second type, where /h/ only occurs syllable-finally or word-finally.

The prediction of diachronic typology, then, is that in the languages of the world, there will be languages in which /h/ occurs as a phoneme and is favored in syllable-final position because it is the reduced form of earlier fricatives such as /f/, /s/, and /x/; in addition, there will be languages in which /h/ is a phoneme, but it does not occur in final position (having been deleted there). To test this hypothesis, I used a sample of seventy-six languages selected to be maximally unrelated, the GRAMCATS sample (Bybee *et al.*, 1994). Forty-five of these languages have /h/ as a phoneme. The hypotheses suggested by the diachronic patterns are borne out in this sample, as shown in (26):

(26)   i. seven languages allow all consonants in coda position, including /h/;

   ii. twelve languages allow /h/ but with restrictions on other consonants: four of these allow /h/ but have heavy restrictions on other consonants (Chacobo allows only fricatives, Koho allows no obstruents at all, Palantla Chinantec allows only velars, Slave allows only /ʔ/ and /h/);

   iii. one language allows /h/ in syllable-final position but not in syllable-initial (Palaung);

   iv. twenty-five languages do not have syllable-final /h/ – ten of these have heavy restrictions on coda position but fifteen allow many other consonants in this position.

More languages preclude /h/ from codas than allow it (see [iv]), but a significant minority not only allows /h/ there, but seems to favor it in the sense that it is one of a small set of allowed coda consonants, as shown in (ii).

---

[5] For multiple examples of (24), see Méndez Dosuna (1996).

No implicational universals can be formulated to cover the situation. It cannot be said that /h/ in codas implies the existence of obstruents in codas (as, e.g., nasals do), nor is the reverse true. Similarly, the evidence would not support an OT constraint forbidding /h/ from codas. There are simply no synchronic universals concerning /h/ in the codas of syllables.

The diachronic paths explain the ambiguous distribution of /h/ in the languages of the world. Since /h/ is both the outcome of weakening and often subject to further weakening itself, a cross-linguistic situation is created in which the frequent occurrence of /h/ in syllable-final position is in itself an unstable situation. Thus, the simple synchronic distribution of phonemes does not always serve to indicate the naturalness or unmarkedness of a phoneme in a certain position; rather the diachronic trajectory for each phoneme needs to be taken into account. As in other cases, considering the diachronic paths of change takes us one more step towards an understanding of language universals.

9   Conclusion

In this chapter, I have argued for the necessity of taking diachrony into account in the formulation of language universals. In my view, linguistic theory must look beyond synchronic generalizations about particular language states to the formational mechanisms that bring linguistic structure into being. Language states come about through the complex interplay of processes at work as language is used. To assess the place of language in the context of human cognitive abilities, it is important to note that most of the processes at work as language is used apply to non-linguistic activities as well. Thus automation, habituation, and categorization can be seen to operate in non-linguistic abilities. Language is highly evolved but not totally distinct from other neuromotor and cognitive abilities (Bybee, 1998).

Discovering cross-linguistic similarities and then positing them as inherent to the language acquisition device oversimplifies the way similarities among languages arise. In addition, it obscures the relation between linguistic abilities and more general cognitive abilities. Only by a detailed study of the mechanisms behind the linguistic changes that create grammatical and phonological systems will we be able to discover what is truly universal in human language.

References

Aarsleff, H. 1982. *From Locke to Saussure. Essays on the Study of Language and Intellectual History*. Minneapolis: University of Minnesota Press.

Abbot, M. 1991. Macushi. In C. D. Derbyshire and G. K. Pullum (eds.), *Handbook of Amazonian Languages*, Volume III. Berlin: Mouton de Gruyter. 23–160.

Abdoulaye, M. L. 1992. Aspects of Hausa morphosyntax in Role and Reference grammar. Ph.D. dissertation, University at Buffalo. (Available on RRG website.)

Acero, J. J. 1993. *Lenguaje y filosofía*. Barcelona: Octaedro.

Ackema, P., and M. Schoorlemmer. 1994. The middle construction and the syntax–semantics interface, *Lingua* 93: 59–90.

Ackerman, F., and P. LeSourd. 1997. Toward a lexical representation of phrasal predicates. In A. Alsina, J. Bresan, and P. Sells (eds.), *Complex Predicates*. Stanford: CSLI Publications. 67–106.

Ackerman, F., and G. Webelhuth. 1998. *A Theory of Predicates*. Stanford University: CSLI.

Aikhenvald, A. Y. 2002. Typological parameters for the study of clitics, with special reference to Tariana. In Dixon and Aikhenvald (2002b), 42–78.

Aksu-Koç, A. A., and D. I. Slobin. 1985. Acquisition of Turkish. In Dan Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*. Volume I: *The Data*. Hillsdale, N.J.: Laurence Erlbaum Associates. 3–24.

Alexandrova, G., and O. Arnaudova (eds.). 2001. *The Minimalist Parameter*. Amsterdam: John Benjamins.

Anonymous. 1980. *Ute Reference Grammar*. Ignacio, Colo.: Ute Press.

Arens, H. 1969. *Sprachwissenschaft. Der Gang ihrer Entwicklung von der Antike bis zur Gegenwart*. Munich: Verlag Karl Alber.

Aronoff, M., and S. Sridhar. 1983. Morphological levels in English and Kannada; or atarizing Reagan, *Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax*. Chicago Linguistic Society 19: 3–16.

Ashton, E. 1944. *Swahili Grammar*. London: Longman.

Bach, E., and R. T. Harris (eds.). 1968. *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston.

Baker, M. C. 1985. The Mirror Principle and morphosyntactic explanation, *Linguistic Inquiry* 16: 373–416.

1988. *Incorporation. A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.