Degenerative Phonology
Daniel Silverman

Part 1
Theory

Chapter Three
Origins

The degenerate character of complex adaptive systems—phonology among them—does not arise instantaneously and fully-blown. Rather, just as a degenerate system is subject to evolutionary pressures once it is in place, degeneracy's very emergence is a product of evolutionary pressures as well. Simply stated, degenerate systems evolve from non-degenerate ones. In this chapter, one possible route to this remarkable transformation—that from a simple, non-degenerate "one-to-one" sound communication system, to a complex, degenerate, and asymmetric "many-to-many" linguistic one—is explored. It is proposed that there need be no magic bullet responsible for this evolutionary transformation, genetic, neural, cognitive, social, or otherwise. Rather, the very pressures acting on a simple system may interact in ways that naturally and passively transform a non-degenerate sound communication system into a degenerate one; a system with many-to-many sound-meaning correspondences, and possessed of robustness, evolvability, and complexity.

Of course, the specific proposals herein may be all wrong, but that's not the point for now. Rather, we are simply interested in exploring the mere possibility that slow-going usage-based pressures on our sound communication system, pressures that were simultaneously both unleashed and constrained by interactions among our developing vocal tract physiology, our developing patterns of socialization, and our developing brain and cognitive structures, may have naturally and passively achieved degenerative linguistic status. As we are specifically interested in the mechanics of an evolving degenerative *phonology*, vocal tract physiology will take pride of place in our musings on the topic.

1. One-to-one sound-meaning correspondence

Consider the nascent stages of our sound communication system, one that was likely qualitatively non-distinct from those of lower animals both past and present, in that it likely involved a one-to-one correspondence between sound and meaning.

The first meaning-imbued sounds of our species may have settled towards ones involving a sudden expulsion of air from the mouth due to an oral seal being broken (plosives), followed by vocal fold vibration accompanying the oral opening gesture (vowels). There are articulatory, aerodynamic, acoustic, and auditory reasons for this.

Regarding articulation, a plosive is quite easy to produce in comparison to other gestures that have come to be part of the speech repertoire, gestures that often require extreme muscular and timing precision to achieve their characteristic aerodynamic, acoustic, and auditory traits (Ladefoged and Johnson 2011). Moreover, upon the simple breaking of an oral seal and allowing air to rapidly flow from the lungs and out the mouth, the vocal folds, when properly postured, may readily engage in vibratory activity (Rothenberg 1968).

Aerodynamically, plosion definitionally involves a forceful and energized expulsion of air from the vocal tract, one without undue respiratory effort. As air is the medium of sound transmission, increased airflow allows for more salient and more varied sounds. Also, again, the subsequent wide open vocal tract creates the proper aerodynamic conditions for the vocal folds to vibrate, thus affording a salient realization to both pitch and resonance distinctions in the acoustic signal.

Acoustically, plosion produces a speech signal of comparatively heightened energy, one in which any number of acoustic modifications might eventually be encoded (among them, those deriving from laterality, labiality, palatality, rhoticity, and vocal fold spreading or constriction). In this context then (that is, immediately following plosion), all these modifications are thus saliently distinct from one another in terms of their spectral characteristics.

Regarding audition, the mammalian auditory nerve is especially responsive to sudden increases in acoustic energy (Delgutte 1982, Tyler, Summerfield, Wood, and Fernandez 1982); a quick reaction to the sudden breaking of silence provides obvious survival advantages in predation situations. The nascent speech code would likely exploit this property from the outset, as all linguistic systems do to this very day (Bladon 1986). And although the release of a nasalized oral closure (that is, a nasal stop) might be no less articulatorily natural than a non-nasalized one, the auditory benefits to keeping the nasal passage shut suggests that our earliest speech sounds may have consisted of orally-channeled air, rather than nasally-channeled air.

In addition to the intrinsic advantages of this most basic of phonetic events, as the vocal tract co-evolved towards its modern incarnation, the location of the oral seal may readily be changed. The seal may be labial, but also, the flexibility of the tongue allows both its front to form a seal at the alveolar ridge, and its back to form a seal at the velum. The perceptual product of these distinct closure locations is three easily-distinguished speech events of exceptionally short duration. This tripartite perceptual distinction establishes the conditions for different acoustic signals to encode different meanings; we might imagine an early stage during which these three closure postures were in place, coordinated with largely undifferentiated qualities to their opening postures, perhaps resulting in three phonetic events that might be recruited toward communicative ends, roughly, **pu**, **ti**, **ka**, each of these phonetic primitives corresponding to single element of meaning, say,“Run!” (**pu**),“Kill!/Eat!” (**ti**), “Sex!”(**ka**). At this stage then, there is a simple one-to-one sound-meaning correspondence, the sort of system that is characteristic of perhaps all our planet's non-human species that engage in sound communication.

2. Many-to-one sound-meaning correspondence

Every speech act is inevitably different from every other: one **pu** is different from the next, even when uttered in extremely comparable real-world circumstances. Indeed, even when their real-world contexts differ such that systematic phonetic differences emerge, these instance-to-instance differences would be unlikely to change core meaning: **pu** may be rendered more excitedly when a prowling lion is spotted as opposed to a lounging one, but even among lower animals, core meaning appears to remain stable despite such real-world context-dependent variation (Seyfarth and Cheney 1992). The crucial factor that renders such *phonetic variation* qualitatively distinct from *phonological alternation* is that the former are context-free in terms of their phonetic surroundings (however conditioned to real-world context they might be), whereas the latter are context-sensitive in this regard, and thus are subject to systematic, context-dependent alternations such that they quickly come to participate in many-to-one sound-meaning correspondences, as we'll now see.

Consider the phonetic consequences of producing two of our meaning-imbued sounds, our “proto-morphemes”, in quick succession. There is any number of ways that such complexity might develop.

For example, two-sound sequences may represent an assemblage of a topic-comment-like element, say **pu-** followed by **-ti** ("Run! Kill!/Eat!") or **ti-** followed by **-pu** ("Kill!/Eat! Run!"), either of which might convey a passive predation warning ("Run if you don't want to get killed and eaten (by that animal)!") or an active predation warning ("Run to kill and eat (that animal)!") (When morphs appear in isolation, the en dash, again, is *not* intended to represent a so-called "morpheme boundary"; rather it is intended to connote any and all the bonded material that varies as a consequence of any additional morphemes' phonetic shapes; it is a variable). Among other possibilities, two sounds may be strung together to name more objects, in a nascent form of noun-noun compounding. Both of these structure-building strategies are present in virtually all languages, of course, but while we will return to the increased semantic complexity that results from such groupings of sounds, for now, consider their phonetic complexities, complexities that culminate in a form-function correspondence that is many-to-one in character.

Indeed, from the moment that a juxtaposition of two sounds is regularly produced, the system achieves this "many-to-one" status. Here's why: as discussed in Chapter One, when one sound is juxtaposed to another, each of the sounds undergoes a systematic change in its phonetic character. Take **pu-** followed by **-ti** ("Run! Kill!/Eat!") as an example. Here, the first sound is systematically modified by the immediate succession of the second, and likewise, the second sound is systematically modified by the immediate precedence of the first. Since the vocal tract posture that accompanies one sound cannot instantaneously transform into the posture that accompanies another sound, the postures affect each other, and the acoustic signal follows suit (Öhman 1966): (1) the first vowel is affected by both the second vowel and the intervening stop in terms of its offset transitions; (2) the intervening stop is affected by the preceding vowel in terms of its onset transitions and its release burst; (3) the second vowel is affected by the first vowel in the form of its onset transitions: **pu-ti**. Thus these two elements' juxtaposition thus results in a temporal span of overlap—a *bond*—that provides phonetic (hence oftentimes semantic) information about *both* sounds.

So, whereas until this time there had been a one-to-one sound-meaning correspondence, now—instantly and irrevocably—this correspondence is sabotaged: the juxtaposition of one sound to another thus opens the floodgates to a many-to-one sound-meaning correspondence.

At these nascent stages then, as sound complexes are repeated and repeated in their appropriate real-world contexts, *new* sounds inevitably arise. This is certainly true of vowels when they come to immediately precede stops, but for now, consider the stops themselves. While constant repetition of juxtaposed sounds in appropriate situations may serve to reinforce their *semantic constancy*, it is their very repetition that induces their *phonetic change* (Kruszewski 1883). For example, the medial closure in our **pu-ti** example may eventually undergo a process of voicing, becoming **pu-di**, intervocalic voicing being a very natural phonetic development (Rothenberg 1968). At this point, both **ti-** and **-di** correspond to a single meaning. This systematic change in sound does *not* expand the inventory of meanings, but it *does* expand the inventory of motor routines put in service to encoding this meaning. Examples of intervocalic voicing are ubiquitous. In Southern Italian for example, the sound pattern possesses the partial voicing of stops when placed in intervocalic contexts, thus **parte** ("part") - **di-b̥arte** ("of a part"), **terra** ("land") - **la-d̥erra** ("the land"), **karne** ("meat") - **di-g̊arne** ("of meat") (Gurevich 2004).

This sort of simple and natural sound change may set in motion a massive increase the sound system's complexity. Indeed, with a larger and larger garrison of *phonetic* elements to deploy, a huge expansion of the *semantic* inventory becomes possible as well, one able to meet the needs of our species' increasingly sophisticated cognitive and social structures: distinct sounds that have heretofore corresponded to a single meaning may now unhinge themselves from their predictable contexts, to be cycled and recycled in ever-increasing and unpredictable ways. For example, newly-voiced medial stops may now appear in first position, for example, **di-bu**, where **di** is an old phonetic element that has been recruited to perform a new semantic role.

Now, a hitch: when this new **di-** of **di-bu** is placed in second position (for example, **ka-di**), it might be pronounced comparably to the closure voicing that had earlier been added to **-ti** in this context (for example, earlier **bu-ti**, now **bu-di**). We thus might fear—during one brief moment of evolutionary panic—that two different semantic elements may now be cued by the same phonetic elements in similar or even identical contexts. That is, we may be moving towards a situation in which we may have **bu-di** ("Run! Kill!/Eat!") where **-di** means one thing ("Kill!/Eat!", but also **bu-di** in which **-di** means something else (say "Avoid predator!"). At this point, a single phonetic form in a single context may perform a dual semantic role: when **-di** finds itself in second position, it is rendered identical to *another* **-di**, that which alternates with **ti-**: running to kill and eat is very different indeed from running away to stay alive!

But such a situation is unlikely to come to pass, especially in intervocalic contexts: as a consequence of the acoustically informative context in which they reside, intervocalic consonants very rarely alternate such that are rendered non-distinct from one another, and so induced homophony is almost certain to be avoided (Gurevich 2004).

Indeed, if many sounds each came to correspond to more than one meaning, listener confusion and communicative failure may result (Martinet 1952, Labov 1994, Silverman 2012). To defeat the pervasiveness of this potentially counter-functional development, the **di-** of **di-bu** may passively undergo another change when found in second position. Since "old" **ti-** now alternates with **-di** when placed between vowels, "new" **di-** may spirantize in this same context, perhaps culminating in **-zi** (or maybe **-ði**). Spirantization of intervocalic voiced stops is likely to take hold *exactly because* of its function-positive consequences: creeping phonetic patterns that eschew undue listener confusion are likely to be replicated and conventionalized. In short, successful speech propagates.

Gurevich (2004) emphasizes how common it is for stops to spirantize intervocalically (typically, though not always, in functional response to the intervocalic voicing of voiceless stops), and how, in 95% of the cases she documents (specifically in seventy-two of the seventy-six cases found in the 230 languages she investigates) the pattern cannot induce homophony (though not all her cases of spirantization involve solely the intervocalic context).

So, we now have **di-** alternating with **-zi**, both meaning one thing ("Defend territory!", and, recall, we have **ti-** alternating with **-di**, both meaning another ("Kill!/Eat!"). The co-evolution of these many-to-one relationships between sound and meaning results in many meaningful elements of the speech signal possessing both systematic phonetic variation and semantic stability across varied contexts. Now, in turn, this new phonetic event **zi** may unhinge itself from its context and be deployed to signal new meanings.

Further developments: maintaining voicing in utterance-initial position is aerodynamically unnatural, oftentimes involving an actively expanded pharynx and a lowered larynx (Rothenberg 1968). Consequently, newly-evolved **bu- di- ga-** might gradually lose this voicing, thus running the risk of sounding the same as **pu- ti- ka-**. If this natural tendency begins to take hold, then those spontaneous productions of *original* **pu- ti- ka-** that possess a slight delay in voicing may emerge as new and different sounds—**pʰu- tʰi- kʰa-**—which now, again, may unhinge themselves and encode new meanings, thus allowing them to appear in second position: **-pʰu -tʰi -kʰa**. English may have proceeded on just this path: word-initial position is characterized by a plain – aspirated distinction among its stops, and a voiced – plain distinction in certain non-initial contexts.

Another possibility: the pitch-lowering effect that naturally accompanies voiced stops may, over time, migrate to pervade the following vowel, coming to replace closure voicing itself, and so becoming a tone distinction that the language may now recycle: **bu- di- ga-** as distinct from **pu- ti- ka-** may yield to **pù- tì- kà-** as distinct from **pú- tí- ká-**. In Northern Kammu, for example, a historic voiceless-voiced stop distinction has evolved into a high-tone – low-tone distinction. Thus, where more the conservative eastern dialect has **taːŋ** (“pack”), and **daːŋ** (“lizard”), the northern dialect has **táːŋ** and **tàːŋ**, respectively (Svantesson and House 2006).

Alternatively again, our phonetically “difficult” initial voiced stops may evolve to be accompanied by velic venting during their oral closures—a tried and true development that often passively evolves—thus again maintaining their phonetic distinctness from **pu- ti- ka-**: **mbu- ndi- ŋga-**. As expected now, these prenasalized forms may unhinge and recombine as **-mbu -ndi -ŋga**, thus opening the gates to phonotactic complexity, say, **kã-mbu**, **kã-ndi**, **kã-ŋga**, and of course, creating more fodder for an expanding inventory of sounds and an expanding inventory of meanings. For example, Flemming (2002) observes that prenasalization of voiced stops may evolve in word-initial position—a context in which such stops are necessarily in contrast with voiceless stops—in Guarani, Barasano, and Rotokas (though this list may be extended with ease), but has not been found to develop in intervocalic stops, a context, recall, in which voicing is easily maintained.

Clearly and emphatically, all these new wrinkles are found time and again on the immortal face of language structure, both as (diachronic) changes, and hence, virtually necessarily, as (synchronic) alternations, such that the system has now passively and naturally evolved from a simple one involving a one-to-one correspondence between form and function to a complex one involving a many-to-one correspondence between form and function.

3. Many-to-many sound-meaning correspondence

The proposed system has now evolved to a stage in which heterophony is ubiquitous. But a many-to-many correspondence between sound and meaning (which includes induced homophony) has thus far been staved off. How then might such a system achieve “many-to-many” status? A few plausible scenarios immediately present themselves, one of which we consider now.

Recall that **bu- di- ga-** are at risk of losing their closure voicing, thus becoming homophonous with original **pu- ti- ka-**. Recall additionally that one route to heterophony maintenance here involves velic venting during oral closure, culminating in **mbu- ndi- ŋga-**, which may unhinge and recombine as **-mbu -ndi -ŋga**. The resultant structures—for example, **kã-mbu**, **kã-ndi**, **kã-ŋga**—may now be snapped at new joints, resulting in new phonetic elements that might acquire unique meanings: **kãm**, **kãn**, and **kãŋ** may now join the repertoire of phonetic/semantic (morphemic) primitives. Indeed, the location of this snap is especially likely, since **bu**, **di**, and **ga** are already part of the sound-and-meaning inventory, and thus their remainders—**kãm**, **kãn**, and **kãŋ**—emerge in high relief as likely candidates for phonetic/semantic deployment.

Now, when **kãn-**, for example, combines with forms like **-bu**, **-di**, and **-ga**, the nasalized alveolar closure is quite susceptible to assimilation (more so, for complex articulatory and acoustic reasons, than are nasalized velar and especially labial closures): **kãm-bu**, **kãn-di**, **kãŋ-ga**. Indeed, nasal assimilation is perhaps the most frequently encountered alternation in the world's languages (Nathan 2008). Here, three phonetically distinct forms (**kãm**- **kãn**- **kãŋ**-) now also correspond to a single meaning, but two of these three forms (**kãm**- **kãŋ**-) might *also* correspond to *other* meanings. Induced homophony has now evolved, and, coupled with the heterophonic alternations already in place, a many-to-many sound-meaning correspondence emerges. The system is now in a state of degeneracy.

4. Asymmetric many-to-many sound-meaning correspondence

Nasal assimilation is especially likely if certain phonetic and semantic conditions are met. As noted, *phonetically*, when an oral closure is not immediately followed by an oral opening—and unlike such stops in intervocalic contexts, as just considered—important *release* cues that might otherwise signal its accompanying oral posture become susceptible to loss. Instead, the oral posture of the following gesture—one that is indeed followed by an oral opening—comes to expand its bond to pervade the nasal murmur itself. Meanwhile, *semantically*, assimilation is more likely to conventionalize if the resulting phonetic string is *uniquely* paired with a semantic primitive, that is, if homophony and listener confusion is not induced.

Nonetheless, this natural assimilative tendency may indeed take hold—thus on occasion inducing homophony—perhaps particularly if the resulting homophone is either very *frequently* deployed (thus increasing its predictability for listeners) or very *in*frequently deployed (thus easing listeners' lexical search). So, in cases when **kãn**- might tend to alternate (**kãm-bu**, **kãn-di**, **kãŋ-ga**) such that homophony is induced with semantically distinct **kãm**- and **kãŋ**-, for example, then the alternation is more likely to take hold if, despite this induced homophony, semantic content is transmitted intact due to the overarching prevalence (hence predictability) or rarity (hence perspicuity) of the morphological complex's semantic content.

Of course, homophonic forms will necessarily be rather few and far between, since an excess of such forms would stymie communicative success, and thus not contribute to the overall structural conventions of the emerging system. Recall: if the same phonetic forms were deployed to both attack defend against a predator, survival of the communicative elements—and, in the case at hand, even the agents who deploy them—would become jeopardized. Developments that enhance the robustness, complexity, and evolvability of the system are selected. Those that don't are not. Homophonic forms that induce listener confusion unlikely to be conventionalized for exactly these reasons.

It is now clear that the bonding which inevitably results from the mere juxtaposition of two simple sounds triggers remarkable growth and complexity of both the phonetic and the semantic inventories. The inevitable consequences of bonding produces both one-to-many and many-to-one sound-meaning correspondences (heterophony and induced homophony, respectively). Moreover, natural, passive, usage-based pressures are in place to ensure that, while heterophony may proceed virtually unchecked, induced homophony remains limited.

The product of this evolutionary trajectory is a degenerate system evincing an asymmetrical many-to-many sound-meaning correspondence.

5. Post-compositionality

Repeated usage of the highly bonded yet still compositionally transparent two-element structures does not only induce the sorts of phonetic changes just explored, but may actually trigger the loss of compositionality itself, resulting in even more complex phonetic spans that come to correspond to semantic primitives. For example, compositional **pu-ti** possesses a meaning that is transparently built from **pu-** and **-ti**. But through its constant use and re-use, in addition to its phonetic changes, it may lose its link to its semantic origins, and thus become stranded as a semantic primitive (Kruszewski 1883); that is, it loses its compositionality, becoming *post-compositional*, or lexicalized, in standard parlance. The now-opaque form (perhaps **puti**, or **pudi** or, **p^huzi**, or **pútì**, for example) becomes a single phonetic form that correlates with a single semantic function, thus embodying a counter-pressure back towards a one-to-one sound-meaning correspondence, even as the system becomes increasingly phonetically complex. Kruszewski provides an example of lexicalization from Latin: **komput-are** (‘’to calculate’’) has achieved post-compositionality as French **kɔ̃te** (to recount), but Latin **refik-ere** (‘’to make again’’) retains its morphological structure in French **ʀə-fɛʀ**.

This tug-of-war between compositionality and post-compositionality thus induces a lengthening of our meaning-impregnated sounds. Whereas earlier, the bonding of one form with another involved only two mouth-opening gestures (of increasingly varied forms), now such juxtapositions may involve three or four mouth-opening gestures, for example, **pu-tika**, **puti-kati**, etc.

Again, e*ach and every one* of these hypothetical developments is not merely a proposed characteristic of the nascent degenerate system. Rather, they are all encountered over and over again in the history of language change. This is not a coincidence. Modern-day pressures on sound patterning are not merely characteristic of the modern-day morpho-phonological system. Rather, they may have been in place long before the system came into existence, acting as a driving and inertial pressure on the very development of the system itself. Natural, systemic, phonetic changes are not merely a *result* of degeneracy. Rather, they are a very *cause* of degeneracy.

6. Composition signals

Although degeneracy had now been achieved qualitatively, still, some systems may be more degenerate than others: there are now pressures that *inhibit* the quantitative growth of degeneracy (manifested as a pressure towards post-compositionality) and pressures that *promote* the quantitative growth of degeneracy (manifested as a pressure toward compositionality).

Consider first a passive *resistance* to the quantitative growth of degeneracy. We have been assuming that context-induced changes to phonetic primitives inevitably trigger their “unhinging”, such that they may come to be assigned additional meanings, and thus come to freely combine in new ways (recall, if **pu-ti** becomes **pu-di**, this new sound involving closure voicing—**-di**—may now be assigned an additional meaning, thus freeing itself from the shackles of its context, allowing for **di-**). Still, if more and more phonetic elements combine into wholly unconstrained sequences, a genuinely damaging ambiguity-of-meaning may result, in the form of an excess of induced homophony. For example, the string **putika** may be ambiguous between compositional **pu-tika** and compositional **puti-ka**.

As a natural consequence of morphs' adaptation to the different contexts in which they find themselves, they may be subject to a passive curtailment in their distribution such that certain sounds are only found in certain contexts. In addition to enhancing and clarifying each morph's phonetic distinctness in terms of its *paradigmatic* patterning, these context-dependent adaptations naturally and passively enhance and clarify each morph's *syntagmatic* patterning as well.

For example, recall that the system may naturally achieve a state in which voiceless stops are limited to sound-*initial* position, and voiced stops are limited to sound-*medial* position, thus **pu-tiga** and **pudi-ka**. Distinctions in stop voicing now act to cue the compositionality of the forms: encountering cues to a voiceless stop in the speech stream confirms that a new semantic primitive is beginning ("Lo! New semantic content afore!"), while encountering cues to a voiced stop signals a continuation of the current semantic element ("Steady as she goes! No new semantic element on the horizon!"). That is, natural phonetic developments may be further harnessed, or *exapted*, to perform new functional roles. Every language passively evolves such patterns, which sometimes go by the name of "boundary signals" (Trubetzkoy 1939). Herein though—in order to resist the temptation to reify the misleading notion of "boundary"—we refer to these syntagmatic cues as *composition signals*.

Heterophony and clarity of syntagmatic structure is thus maintained in a decidedly passive way, simply because those speech acts that are not semantically ambiguous are likely to be the very ones that are communicated successfully, hence imitated and conventionalized. Indeed, in most languages, the phonetic properties of word-initial stops are different from these properties in word-medial position, thus serving this composition-signaling function. For example, Trubetzkoy (1939) reports that in Barra Gaelic, aspirated stops are found exclusively in word-initial position. Consequently, aspiration serves a dual function here: (1) it provides a salient distinction with the plain stops that are contrastive in word-initial position, and (2) it serves as a salient composition signal: an aspirated stop means a new word has begun. Thus, both paradigmatic and now syntagmatic patterning are passively shaped and cued by natural phonetic developments.

Still, even in the absence of these particular sorts of composition signals, most languages have extremely reliable cues to composition in the form of prominence or *stress*. Let's return to our phonetic span **putika**. Even in the absence of medial closure voicing, clarity of compositional structure may be conveyed by stress, say **ˈpu-ˌtika** or **ˈputi-ˌka**; one stress per semantic primitive. These stress distinctions serve to structurally—and, in most cases, semantically—disambiguate phonetic spans that might otherwise sound the same. Reflecting its proposed origins as an aid in disambiguating these early two-sound structures, stress typically involves a binary iambic or trochaic rhythmic pattern at word edges, often iteratively applied in accommodation to the inevitably increased length of meaningful elements of the speech stream we are now considering, that is, words and phrases (Hayes 1995). The role of stress as a binary *phonetic* structure that may have originally cued a binary *semantic* structure thus persists, in remarkably comparable function and form, up to the present day.

Our nascent speech code now possesses both the *more*-assimilated spans characteristic of *frequently* juxtaposed semantic elements, and the *less*-assimilated spans characteristic of *rarely*-juxtaposed semantic elements. Composition signals—in the form of *strong* bonding among frequently juxtaposed elements, and in the form of *weak* bonding among less-frequently juxtaposed elements—may now be seen to induce the emergence of so-called "words"; morphological complexes that are frequently cycled and recycled as necessary for communicative success.

7. Constituency

Composition signals are not ubiquitous. In the absence of such signals, a genuine counter-functional ambiguity-of-meaning will, on occasion, be present in the speech code. Remarkably though, it may be the very ambiguity of our increasingly complex phonetic spans that establishes the conditions for hierarchical morpho-syntactic structure to arise: semantic ambiguity of structural origin feeds a hierarchical constituent-structural analysis.

Consider our **putika** case again (assuming for now the absence of any morpho-syntactic composition-signaling phonetic content). At these early stages, recall that at least two structures and meanings may be paired with this single phonetic span: **pu-tika** and **puti-ka**. In most cases, real-world knowledge will serve a disambiguating function, but once in a while, genuine ambiguity necessitates a deeper structural analysis by listeners ("Is it **pu-tika** or **puti-ka**?"). But note that the very moment listeners consider competing structures and their associated meanings, they are engaging in constituent analysis: the potential for hierarchically-structured morpho-syntactic strings suddenly becomes a reality.

The semantic ambiguity exemplified by **pu-tika** versus **puti-ka** is qualitatively distinct from what we have considered thus far, as it is an ambiguity rooted in *structure*, not an ambiguity rooted in the mere phonetic identity of semantically distinct primitives (homophonic morphs). Such phonetic spans' semantic ambiguity triggers their deeper structural analysis. Listeners' rising to the challenge of structural ambiguity, then, opens the gateway to morpho-syntactic hierarchical constituent structure, by requiring these listeners to perform deeper structural analyses of received phonetic signals than had been heretofore required. The ambiguous affiliation of the bonded material thus opens the gates to hierarchical constituent structure.

Of course, these multiple interpretations of particular phonetic strings should be few and far between, since most phonetic events possess (1) phonetic cues, (2) semantic cues, and (3) pragmatic cues, to the intended structure and meaning of the span. Consequently, and most interestingly, it is exactly those rarely-encountered ambiguous forms that might trigger the emergence of a hierarchical and recursive organization.

8. Recursion

Now consider a longer string that is ambiguous, for example, **putikakatipu**. This string might be intended by the speaker as, say, **putika-katipu**, and yet is open to a number of interpretations by the listener. For example, imagine the ambiguous affiliation of its middle span, very roughly **kaka**: both **putikaka-tipu** and **puti-kakatipu**, may be perceived, assuming each of these makes sense to the listener. So far, this is exactly the scenario just considered with respect to **putika**.

Clearly though, in comparison to **putika**, this longer string is impregnable with many more structures and meanings. Consider, for example, **[[pu-ti]-kaka]-tipu**, or **puti-[kaka-[ti-pu]]**, or **[[puti]-ka]-[[kati]-pu]**, (where some brief spans here actually bear the mark of *three* morphemes, not the two that typographical limitations suggest; double under- and overscoring are employed in an effort to enhance compositional clarity). Perhaps more than one of these distinct parses might be sensibly interpretable by listeners under the appropriate real-world conditions, even if the speaker intends a "flat" non-hierarchical binary or even unary structure. Again, it is the semantic ambiguity of the string that triggers its deeper structural analysis, an analysis that quickly culminates in both hierarchical and now *recursive* structures, when embedding involves elements of the same type. Indeed, recursion is considered by some to be a primary characteristic of grammar (Hauser, Chomsky, and Fitch 2002).

In sum, the phonetic product of two juxtaposed sounds of increased length may lack semantic clarity due to an ambiguous affiliation of its bonded span. The resulting string is thus ambiguous between (at least) two different structures, thus triggering deeper analyses on the part of listeners, culminating in these sounds' hierarchical structuring, and further, opening the floodgates to recursion.

In short, induced homophony of structural origin may have triggered the emergence of morpho-syntactic hierachical complexity.

9. Subsumption

Recall that degenerate systems—embodied as the many-to-one and one-to-many relations between form and function—possess elements that are at once (1) sufficiently impervious to insult such that they remain vital to the proper functioning of the system as a whole (culminating in the system's *robustness*), (2) sufficiently variable such that they might adapt to new conditions coming to act on their form (culminating in the system's *evolvability*), and (3) sufficiently interactive such that they enter into a hierarchical organization (culminating in the system's *complexity*).

The degenerate *system* possesses these qualities, but any individual *component* of the system may nonetheless be susceptible to weakening and even loss. In a degenerative phonology, for example, usage-based phonetic pressures may induce an eventual withering away of a given element. Despite its phonetic demise though, this withered element's function may be subsumed by a more complex structure (a fusion, a phrase, or a construction, for example), and thus the system's robustness, evolvability, and complexity are maintained. In degenerative phonology, the catalyst of such a *subsumption* is, as with so many other aspects of the system, the bond.

Take a schematic example. Perhaps the final phonetic span of a morphological complex like **kati-pu**—that is, **pu**—as a consequence of its perceptually inauspicious word-final context, begins to weaken towards zero. But given that this phonetic span plays an important role in signaling semantic content, its eventual *phonetic* demise is unlikely to be accompanied by its *semantic* demise. Instead, the early portion of its bond is likely to take up the functional slack: the minor labiality that had heretofore appeared on the preceding vowel may grow in its *formal* perspicuousness exactly because of its growing *functional* importance, perhaps culminating in a front rounded vowel that now allows for the entire loss of the increasingly extraneous phonetic material that follows: **kati-pu** evolves into **katy-pu** evolves into **katy-p** evolves into **katy**. The relevant domain now possesses a fully *fused* element, the bond having taken over the full brunt of encoding the semantic content of what were previously separate elements.

Such subsumptions, note, are only possible when bonding is present. The function of one element or structure may be fully overtaken by another element or structure only if there is a historic period of a multiplicity of phonetic cueing such that some cues co-vary in a trading relationship. In our example case, historically intermediate vowel harmony is fully subsumed by fusion. Germanic umlaut provides us with a well-known example. Simplifying considerably, early **muːs-i**, ("mice") yielded intermediate **myːs-i**, in which the suffix has fully bonded with root content, culminating in this first vowel's fronting. A further development involved the withering of the second vowel itself, thus **myːs**. (After de-rounding and the Great Vowel Shift, we've got **maɪs**.) In short, the suffix and its bond to the root were subsumed by umlaut, culminating in a fused span conveying the semantic character of *both* earlier forms.

Subsumptions do not require the involvement of a grammatical category, however: lexical categories too may undergo subsumption. For example, the English portmanteau "smog" (**smag**), historically derives from "smokey fog", though is now fully lexicalized, referring to any pollutant that limits sight distance under particular weather conditions. More to the point, any of **smag**, **sm-ag**, or **smag** may be appropriate morpho-phonetic transcriptions, depending on the explicit knowledge any individual speaker acquires about the form's etymology. Indeed, these three alternatives are ordered here in a way that likely mirrors the form's ontogenetic evolution.

Phonetics and semantics are always talking to each other, and one such sort of dialogue may result in subsumption. The existence of subsumption thus does not weaken the degenerative phonology proposal. Rather, it bolsters it: due to the inevitable interaction among a complex hierarchical system's elements, the emergent variation of its forms provides the very fodder for both the system's and its elements' maintained functionality. Conflicting pressures yield system-internal modifications that may take hold exactly because of their functional efficacy: robustness, evolvability, and complexity become more intertwined as degeneracy proceeds.

In short, these properties of robustness, evolvability, and complexity are both characteristic of, and dependent on, bonding, hence allowing for subsumption. One thus might propose that the traditional typology of morphological systems—concatenative, fusional, analytic, polysynthetic, among others—is better viewed in gradient terms, with different languages plotted at different points on a sliding scale of bonding, as proposed, for example, by Simpson (2009).

10. Productivity

As users master their ambient system, the ability to produce novel forms naturally emerges. Recall our case of nasal assimilation: **kãm-bu**, **kãn-di**, **kãŋ-ga**. Here, three phonetically distinct forms (**kãm-**, **kãn-**, **kãŋ-**) correspond to a single meaning: three morphs; one morpheme. Which morph, then, might a speaker deploy in novel contexts, say, preceding newly-learned **pʰu**, **tʰi**, and **kʰa**? While the answer is obvious (**kãm-pʰu**, **kãn-tʰi**, **kãŋ-kʰa**); the motivation is perhaps somewhat less so.

Recall that the "morpheme boundary" symbol ("-") represents a *variable*, such that each of **kãm-**, **kãn-**, and **kãŋ-**, and also, each of **-bu**, **-di**, and **-ga**, involves phonetic material that breaches its relevant typographic edge: bonded material simultaneously contains—and conveys—phonetic information associated with multiple morphemes, and is best conceptualized as an intrinsic part of the morpheme with which it is bonded. Users know that **kãm-** is deployed only when **-bu** follows, that **kãn-** is deployed only when **-du** follows, and that **kãŋ-** is deployed only when **-gu** follows such that one intervocalic oral closure is employed, though its location is different depending in the phonetic properties of the second element.

At this stage, users have *no experience* with complexes possessing multiple intervocalic closures—that is, with heterorganic nasal-stop clusters—and are hence extremely unlikely to spontaneously produce such forms. Consequently, based on their motoric experience with "old" forms **kãm-bu**, **kãn-di**, **kãŋ-ga**, *morph selection is virtually automatized even for novel constructions*, and thus **kãm-pʰu**, **kãn-tʰi**, **kãŋ-kʰa**, are effortlessly selected, produced, and established as pronunciation norms. Due to the bond, and the phonetic material that is shared by multiple morphemes, morph selection even in novel contexts becomes trivial: motor routines imposed on the morphological string are highly unlikely to stray from those that have already been internalized and routinized through experience.

But there are also semantic factors involved in morph selection. An example from an "r-dropping" dialect of English makes this clear. Consider the noun **wɪn** ("win") that is to be pluralized for the first time. Based on experience with pluralization, there are three candidates: **wɪn-z̥**, **wɪn-s**, and **wɪn-əz̥**, all of which involve motoric activities that, let's suppose, have long been routinized components of a speaker's repertoire: “winds”, “wince”, “winners”. In this case, if morph selection were solely dependent on motor experience, our innovator might be at a loss to conclude which morph to select. Instead though, morph selection is influenced by *semantic* as well as motoric experience: only **wɪn-z̥** conforms to patterns already in place. After all, users have ample experience with the bond—roughly **ɪn-z̥**—when employing plurals, and have absolutely no experience with the other two hypothetical bonds—roughly, **ɪn-s** and **ɪn-əz̥**—in this same morpho-semantic context.

In short, experience with both the phonetic and the semantic properties of bonds may play vital roles in innovative morph selection, that is, in productivity.

11. Contingency

The multi-dimensionality of any phonetic signal and its perception—each and every one involving complex interactions between the precisely controlled aerodynamic and articulatory configurations produced by speakers, and the consequent acoustic and auditory complexities affecting listeners' perceptions, coupled with the cognitive, the pragmatic, and the inevitable “top-down” factors that additionally influence phonetic forms and their semantic interpretations—result in the plain certainty that even extremely slight variations may be reproduced and iterated, culminating in long term effects that change the overall shape of system. In phonology, an *un*limited number of phonetic states is possible within a *de*limited phonetic space. Thus, despite the current proposal that linguistic sound systems may derive from a single source, there is a virtual infinity of *contingencies* serving to influence any given system's future state.

We have already proposed some of these contingencies in action: word-initial voiceless stops *may* come to aspirate, *may* come to induce vocalic tone, *may* come to pre-nasalize. Speakers on one side of the hill may embark on one trajectory of change, speakers having moved to the other side may send the system somewhere else.

The contingencies inherent to any complex adaptive system are thus subject to both variation and iteration of its affected elements, precluding the possibility of confidently predicting its future state, and, more fundamentally, precluding the possibility of exhaustively characterizing or explaining its form at any given stage of its evolution.

Indeed, it would evince both the height of arrogance and the depth of ignorance to propose that the forms that particular—or certainly that *all*—linguistic systems take might be exhaustively characterized or explained: the myriad phonetic and semantic pressures on its form, the myriad cognitive and pragmatic factors that both speakers and listeners bring to bear on the tasks of language production and language perception, clearly embody the emphatically *contingent* nature of any system's shape at any given point on its diachronic trajectory.

11. Summary

It may or may not be relevant that the acquisition of phonology by children proceeds on a trajectory that reasonably hugs the levels of complexity proposed herein for the origins of grammar itself, just as it may or may not be relevant that implicational hierarchies concerning phonotactic complexity, also, fit rather snugly into these proposals. Still, there is no evolutionary-biological privilege bestowed upon the proposed primordial structures that persist into the the present and beyond, just as there is no evolutionary-biological privilege bestowed upon the pentadactyl configuration among our planet’s tetrapods. In both cases, there was merely a sensitivity to an initial complex of conditions that culminated in these features’ prominent role in the evolution of species. Recall: degeneracy is a *consequence* of evolved systemic complexity, not a cause.

The musings on the emergence and maintenance of a degenerative phonology just presented nonetheless demonstrate that the system, unique though it may be in the annals of the known universe, is not “special”: the very same sorts of pressures and principles that affect the emergence and maintenance of other complex adaptive systems are active in the emergence and maintenance of the linguistic one.