Problems & Paradigms

Spoken language achieves robustness and evolvability by exploiting degeneracy and neutrality

Bodo Winter

As with biological systems, spoken languages are strikingly robust against perturbations. This paper shows that languages achieve robustness in a way that is highly similar to many biological systems. For example, speech sounds are encoded via multiple acoustically diverse, temporally distributed and functionally redundant cues, characteristics that bear similarities to what biologists call ''degeneracy''. Speech is furthermore adequately characterized by neutrality, with many different tongue configurations leading to similar acoustic outputs, and different acoustic variants understood as the same by recipients. This highlights the presence of a large neutral network of acoustic neighbors for every speech sound. Such neutrality ensures that a steady backdrop of variation can be maintained without impeding communication, assuring that there is ''fodder'' for subsequent evolution. Thus, studying linguistic robustness is not only important for understanding how linguistic systems maintain their functioning upon the background of noise, but also for understanding the preconditions for language evolution.

Keywords:
degeneracy; evolvability; language evolution; neutrality; robustness

Corresponding author:
Bodo Winter
E-mail: bodo@bodowinter.com

Introduction

When speakers of a language communicate with one another, they frequently do so in noisy and sometimes rapidly changing acoustic environments. Speech is furthermore characterized by a large amount of variation, both within and across speakers [1–3]. One of the reasons why language users can communicate at all, despite such variation and noise, is because they flexibly adapt in perception, making use of the available context to make inferences about the intended meaning of utterances. For example, listeners can mentally ''fill in'' missing sounds by virtue of their knowledge about words [4]. However, besides such listener-based flexibility, the linguistic system itself is structured in a way that achieves robustness to noise and variation. ''Degeneracy'' and ''neutral networks'' – two concepts borrowed from systems biology – are used as conceptual tools to help understand the robustness that is inherent in the structure of the speech signal.

The identification of analogous features between language and biology has a long history (e.g. [5–8]), going all the way back to Darwin, who drew explicit parallels between the family trees of languages and the family trees of species [9]. Multiple fruitful avenues of research sprang from the language/biology analogy, including the application of phylogenetic methods [10] and network approaches [6, 11] to linguistic data. This paper argues that the study of robustness provides a further point of synthesis between biology and linguistics.

Is spoken language robust?

Robustness has been characterized as the ability of a system ''to maintain its functions despite external and internal perturbations'' [12], or as the ''maintenance of some desired system characteristics despite fluctuations in the behavior of its component parts or its environment'' [13]. Before studying how a system achieves robustness, one first needs to establish whether a system is, in fact, robust [14]. In biology, this can be achieved by perturbation experiments, e.g. in gene

knockdown studies. For example, somewhere between 89 and 96% of 16,000 *Caenorhabditis elegans* genes can be subjected to knockdown interference with no measurable phenotypic consequences [15]. Similar studies in yeast [16] and other organisms reveal that interfering with a large number of genes is possible without drastically affecting phenotypic fitness.

Analogous perturbation experiments have been conducted in speech science. For example, speakers are able to deliver accurate renderings of speech even with weights attached to their jaws [17] or pellets to their tongues [18]. They can also talk while biting a block [19], or while a mechanical device pulls the jaw in the opposite direction of intended speech movements [20]. These perturbation experiments reveal what is called "motor equivalence" (e.g. [21]), which refers to the fact that there are multiple production strategies (i.e. different tongue and jaw configurations) that lead to linguistically equivalent acoustic outputs.

In a related class of experiments – auditory feedback experiments – spoken utterances are recorded in real-time and rapidly played back to the speaker in an altered fashion, creating the illusion that one's speech production has changed. These studies have shown that speakers rapidly adapt their speech production strategies to a variety of auditory perturbations (e.g. [22, 23]). On the perception side, understanding spoken language is possible despite many different kinds of acoustic manipulations [24–26]. Taken together, these perturbation studies empirically establish that speech is, in fact, aptly characterized as a robust system.

In biology, another source of evidence for robustness comes from looking at comparable traits in closely related species [14]. For example, there are different developmental pathways that lead to a similarly shaped vulva in *C. elegans* and related species, showing that the same structure and function can be built in different ways [27]. Multiple intermediate developmental stages lead to the same phenotype, effectively buffering vulva development against stochastic variation at the genetic or environmental level [27].

Language is rife with structurally analogous cases. On a broad scale, language diversity itself attests to the robustness of language. There are strikingly different linguistic systems [28], all of which appear to serve similar communicative demands [29]. That the same linguistic functions can be, and are being, fulfilled in different ways is most apparent when looking at comparisons between closely related languages. For example, in the Slavic branch of the Indo-European language family, the languages Polish, Russian, Bulgarian, and Czech show considerable variation in the tongue shapes that articulate the "same" sound [30]. It has been stated that it is impossible to point to a single case where the corresponding sounds of two different languages have exactly the same acoustic and articulatory targets [31]. This means that the same function within a system of oppositions – i.e. differences in sound create differences in meaning, e.g. between *desk* and *disk* – can be realized in multiple equivalent ways.

Such neutrality with respect to communicative function is also evidenced by how linguistic oppositions change over time. For example, voicing contrasts (***time*** vs. ***dime***, ***pear*** vs. ***bear***, etc.) may develop into tonal systems, where pitch

differences come to encode meaning differentiations that were previously signaled by the presence or absence of vocal fold vibration [32]. An illustrative example of this process comes from the Athabaskan language family (spoken in North America). Where the Athabaskan language Chipewyan marks particular words with a low tone, that is, with low or falling pitch, the language Gwich'in marks the same word with a high tone, that is, with high or rising pitch [33]. Thus, two acoustically opposite realizations (low pitch vs. high pitch) serve similar functions within the respective linguistic systems. Because these languages share a common ancestor [33], these two equivalent solutions must have been accessible to language evolution. Thus, with analogy to the "many ways of building a body" [14], we can say that there are "many ways of conveying the same meaning", and over the course of language change, these different linguistically equivalent solutions can develop from the same ancestral system.

Now that we have established that language is characterized by a considerable degree of robustness, we can ask the question: How is this robustness achieved?

Spoken language achieves robustness via degeneracy

One way in which biological systems may achieve robustness is via degeneracy [34–36], which, in systems biology, refers to phenomena where multiple *structurally different* elements perform overlapping function within a system. Crucially, in contrast to everyday usage of the term, "degeneracy" when used in biology has no negative connotations, i.e. it is not taken to imply deterioration or degradation. Degeneracy is merely used as a technical term to contrast with "redundancy of parts". In the latter, *structurally equivalent or repeated* system components realize the same function. In degeneracy, on the other hand, components fulfilling the same function are different from each other, and they may simultaneously perform additional functions in other domains.

An excellent biological example of degeneracy is the human brain [37, 38]: Multiple brain areas are able to perform the same cognitive function, while at the same time, each brain area also serves other functions. The afore-mentioned gene knockout studies have also been interpreted as an instance of the general principle of degeneracy (e.g. [34]), where different genes have overlapping functions and can compensate for each other when knocked out. Thus in biology, the concept "degeneracy" has a positive connotation due to its proposed adaptive value.

As is known by linguists and speech scientists, there are multiple acoustic cues for almost every linguistic entity that needs to be conveyed [39–42]. Take, for example, the words *rupee* and *ruby*, which differ with respect to the so-called "voicing" of the medial consonant – the sound in *rupee* being characterized by linguists as "voiceless", and the one in *ruby* being characterized as "voiced". Voicing distinctions carry a high functional load: that is, they are communicatively important because they distinguish many different words from one another (e.g. *bat* vs. *bad*, *bet* vs. *bed*, *dusk* vs. *tusk*,

Problems & Paradigms

bear vs. *pear*). And, not surprisingly, given the importance of voicing within the English linguistic system, it is degenerately encoded by a multitude of different and partially overlapping cues (see Fig. 1).

When characterizing the voiced/voiceless distinction, linguists have traditionally emphasized what is called "voice onset time" [43]. This acoustic parameter refers to the time it takes for the vocal folds to start vibrating after the consonantal closure is released, usually with a puff of air. Durational differences of this parameter have been shown to be strong determinants of perceiving a consonant as voiced or voiceless. However, a host of other work by speech scientists has revealed numerous other cues for the contrast between "voiced" and "voiceless" stops, including, among others, formants (spectral bands that generally signal vowel distinctions) [44, 45], pitch in the following vowel [46], the duration of the preceding vowel [47], the duration of the consonantal closure [48], as well as loudness differences within the voice onset time [49]. Finally, the voicing distinction is even cued on sounds that are far away from the actual consonant: for example, in the words *le**d*** versus *let*, voicing induces a perceptible acoustic difference in the initial /l/ of the word [50]. Many of these cues serve other functions within the linguistic system: for example, formants not only signal consonantal distinctions, but at the same time they also co-signal vowel distinctions (e.g. [42]). The adaptive significance of having a multiplicity of cues is highlighted by what speech scientists call "trading relations", where changes in a particular acoustic dimension are perceptually compensated when other cues are present [51].

The example of voicing highlights how multiple acoustically diverse cues signal the same linguistic contrast, i.e. the opposition between "voiced" and "voiceless". The distributed nature of this multitude of cues has conceptual parallels to the notion of "distributed robustness" in biology [52]. We may surmise that having the cues spread out over the acoustic

signal makes speech robust against noise that occurs at specific time points. Due to the temporally distributed nature of speech cues, a contrast whose cues are occluded by noise at any one particular time point can still be perceived because of cues at other time points. An illustrative example of this process is provided by linguistic focus, which is used to contrast or highlight certain elements of utterances, usually by an increase in pitch and loudness. However, frequently – and in many languages – focus is additionally marked by a *lowering* of pitch following the focused sentence part, a phenomenon called "post-focus compression" [53]. It has been shown that even when the focused word itself is masked by noise, listeners are able to retrieve the position of focus within a sentence by exploiting the knowledge that speech following focus is compressed [53]. Thus, there is empirical confirmation that the temporal distribution of cues contributes to speech perception in noise.

A particularly striking example of degeneracy in speech comes from cues for word boundaries: Linguistic information is generally "smeared" across the speech signal. For many words and sentences, there are no silent pauses. Most words and sentences follow each other rapidly and are subject to *coarticulation*, which describes the phenomenon that a given speech sound is produced differently depending on the preceding and following sounds. This acoustic smearing makes the perception of word boundaries difficult; and, not surprisingly, segmenting speech is still a technical challenge for many automated speech recognition systems. However, for human perceivers, word boundaries are cued in multifarious ways, including rhythmic patterns [54], predictable patterns of coarticulation [55], statistical distributional information [56], durational differences [57] and sonority [58]. Furthermore, this degeneracy of cues for word boundaries has been demonstrated to have adaptive significance: A neural network can learn to predict word boundaries better when able to use multiple cues [59].
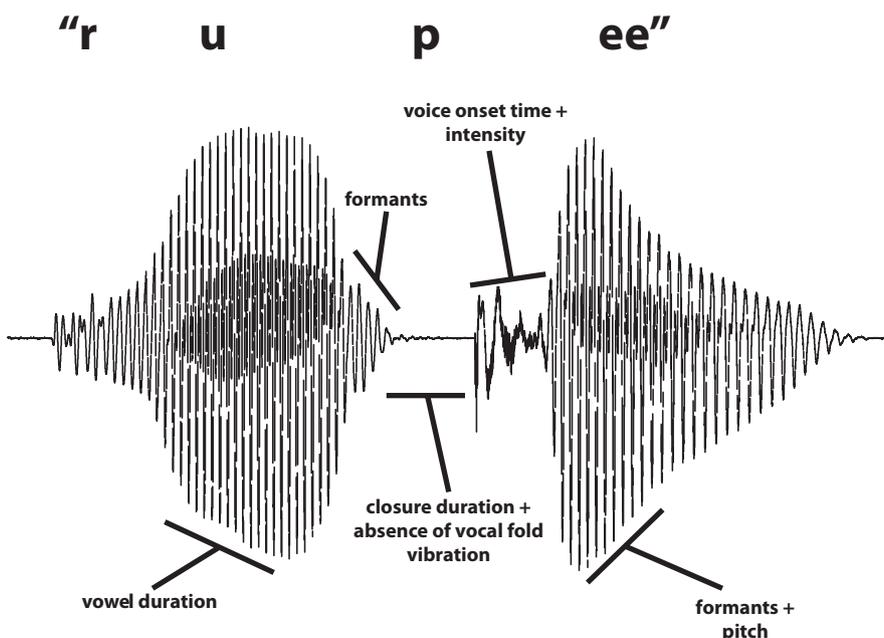

Figure 1. Waveform diagram exemplifying that speech is characterized by distributed robustness, or degeneracy. For the voicing contrast between /p/ and /b/, there are multiple cues that are acoustically diverse and temporally distributed.

Degeneracy is important not only for robustness itself, but also for understanding how spoken languages evolve. The fact that linguistic oppositions are degenerately encoded via multiple acoustically diverse cues means that signaling of an opposition can shift from one cue dimension to another. As a recent example of this, Seoul Korean is reported to be currently undergoing sound change [60, 61]: a consonantal contrast that was previously encoded simultaneously via voice onset time (see explanation above) *and* pitch is relying to an increasing extent on pitch. For the affected sounds, the voice onset time distributions of younger speakers are overlapping nowadays: pitch has come to carry the primary functional load of signaling the contrast. In this example of Korean "tonogenesis", a contrast is preserved, but the underlying acoustic dimension has shifted. However, the fact that the system is able to shift at all from one cue to another without loosing important communicative functions shows that encoding linguistic entities degenerately is an important precondition for contrast-preserving changes.

We may view changes such as the Korean one as a trajectory through a neutral space of communicatively equivalent solutions. In Fig. 2, the horizontal axes mark the two respective cue dimensions. In the Seoul Korean case, these could be voice onset time and pitch. The height of the figure represents intelligibility or communicative fitness. Point A marks the point from which the Seoul Korean system evolved, with strong differences between voiced and voiceless stops for both cue dimensions. The language change that is occurring right now can be characterized as a trajectory to point B, a system that is predominantly cued by pitch. Point C is another possible linguistic system, where the other cue dimension, voice onset time, dominates. Thus, A, B and C lie on a fitness plateau where multiple acoustic realizations exist for the same sound. This shows neutrality with respect to communicative outcomes.

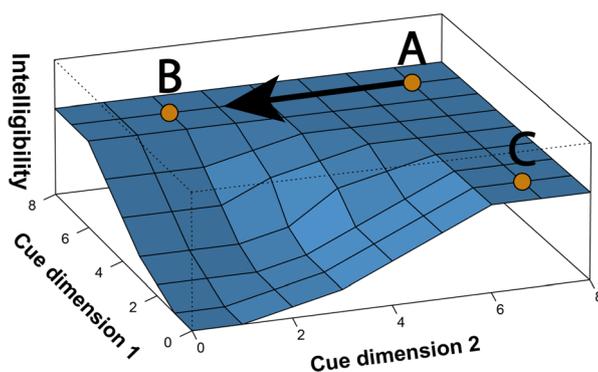It should be emphasized that Fig. 2 is a low-dimensional projection of a much higher-dimensional space. This is because for Seoul Korean – just as for other contrasts and other languages – we know of many more cues that co-signal the same linguistic opposition than just the two cues depicted.

When making the analogy to fitness landscapes in biology, it should also be emphasized that Fig. 2 only depicts one dimension of "phenotypic fitness", namely, the dimension of communicative intelligibility. In biology, researchers have argued against an essentialist notion of neutrality [62]. A mutation, for example, is not neutral with respect to everything. Neutral mutations might have differential fitness if the environment changes, or they might have non-neutral effects on another trait than the one that is being studied in a particular experiment. Similarly, the neutrality depicted in Fig. 2 is only neutrality *with respect to communicating a linguistic message*. However, speakers communicate not just to convey information, but also for other reasons, e.g. to signal social status [63]. Thus, one has to imagine that there are other dimensions, such as the social or expressive value of a linguistic form, that contextualize the low-dimensional space depicted in Fig. 2. A space of solutions that are equivalent from the standpoint of transmitting information can be seen as a prerequisite for these other dimensions of communicative fitness. The mere presence of linguistic variants that are more or less "prestigious" or expressive requires that speakers are able to choose among multiple variants that are able to convey the same linguistic information.

Spoken language achieves robustness via neutrality

The preceding discussion has already highlighted that neutrality – the presence of multiple equivalent solutions – is one important factor in achieving robustness in speech: linguistic variants can be seen as being located in a large neutral space that is spanned by multiple different acoustic cues. Neutrality is also characteristic of two other speech phenomena, depicted in Fig. 3.

The first one, called quantality, refers to the fact that there are regions of motor articulatory space where variations in motor input create little variation in acoustic output (regions I and III in Fig. 3A). In these regions, variation in motor configurations (such as different positions of the tongue) has no audible effect. A classic example of this nonlinear mapping between motor space and acoustic space is the difference between /s/ as in **s**ell and /ʃ/ as in **sh**ell [41]: If one slowly moves one's tongue from /s/ to /ʃ/, there is a relatively sudden transition between the two sounds, with large regions that are equally good renderings of either /s/ or /ʃ/ respectively. Many different speech sounds are characterized by such *quantality* – the characteristic of being *quantal* [39, 64–66].

A second nonlinear transformation is categorical speech perception, where linear acoustic changes are perceived nonlinearly by listeners [67], as following a roughly sigmoidal curve (see Fig. 3B). For example, if listeners are asked to distinguish between the syllables /**b**a/ and /**p**a/ when listening to a linearly stepped continuum between the two sounds, they report to perceive /ba/ for most of the range and

Figure 2. A neutral landscape of possible linguistic systems that all rank equally with respect to "communicative fitness". Each cue alone (*x*- and *y*-axis) achieves similar intelligibility levels when signaling a linguistically relevant contrast. System A is degenerate: a linguistic contrast is encoded by two cues. Within the neutral space, linguistic systems can drift without affecting communicative outcomes. The depicted cue is a low-dimensional projection of a phenomenon with much higher dimensionality, as there are many other cues for a given linguistic contrast.
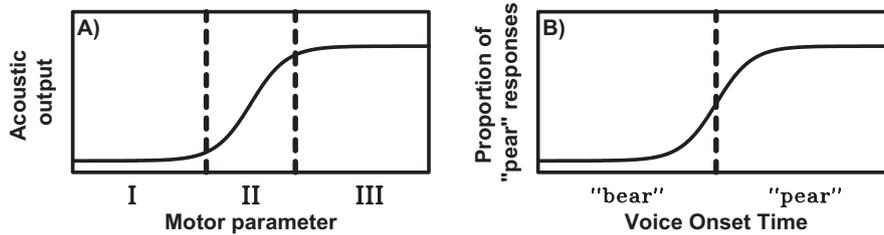
Figure 3. A: Quantality: a nonlinear mapping of motor parameter space (e.g. a particular tongue configuration) to acoustic output. **B**: Categorical perception: a nonlinear mapping of acoustics to perception.

then suddenly switch to perceiving /pa/ at the boundary between the two linguistic categories (cf. [67]). Within the stable regions of Fig. 3B, listeners do not report perceiving differences between different points along the continuum. For many different speech sounds, it has been demonstrated that they are perceived in such a categorical fashion [68, 69].

It is noteworthy that the presence of nonlinear transformations in speech has close parallels with many biological systems, where frequently there are parameters (e.g. enzymatic activity in a metabolic pathway) that can be changed in a "sloppy" fashion, with minimal changes in outcome for many different input configurations [70]. This nonlinearity is intrinsically linked to the notion of robustness, as embodied in characterizations of robustness as "the property of a system to produce relatively invariant output in the presence of perturbation" [71]. In the case of quantality and categorical perception, large variation in input often does not lead to large variation in output, rendering the system more stable.

Quantality is primarily a physiological phenomenon, grounded in the motor-to-acoustic mapping inherent in the physical make-up of the speech apparatus. Notwithstanding the interesting question of where quantality comes from, we may view it as a biological feature of the human vocal tract that a linguistic system can readily exploit to achieve robustness, i.e. by virtue of allocating the sounds of a linguistic system to the quantal regions of the motoracoustic space rather than close to the unstable boundaries of these regions (cf. [39]).

Compared to quantality, which is physiologically grounded, categorical perception is a much more variable perceptual phenomenon: categorical perception boundaries can differ between different languages. However, categorical perception, too, can be seen as something that adult language users are biologically endowed with. Categorical perception has been shown to exist for non-speech sounds as well [72], and it has been shown to be an ability shared with many animals, including, among others, chinchillas [73], budgerigars [74] and crickets [75]. Categorical perception has also been shown to arise early in human ontogeny [76]. Thus, categorical perception is, in many ways, a pre-linguistic nonlinearity that a linguistic system can exploit to achieve robustness of communication.

For both quantality and categorical speech perception, we can say that variation within the neutral regions is

not "visible" to communicative pressures, and a similar argument applies to variation within the neutral regions created through having lots of different acoustic cues. From the perspective of biological systems, such communicative neutrality of different variants of speech sounds means that variation within the neutral space is not selected against. Precisely because variation within the neutral spaces is *not* relevant for the transmission of linguistic information, it persists. Such variability can be called "cryptic", or variability "under the rug" [77] that is hidden from selection.

However, variation that does not matter for communication now may become communicatively relevant later on, for example when the linguistic system changes. That is, the "rug" that previously shielded variation from selective pressures may eventually shift or shrink [77]: previously unseen variation can now be selected for. It should be emphasized that this idea goes beyond the old and well-established principle that variation is a prerequisite for evolutionary change [1, 78]: neutrality (achieved via degeneracy, quantality, and categorical perception) allows the maintenance of variation as fodder for future evolution *while at the same time* supporting the system's robustness. Neutrality means that the system *affords* more variation and increases the evolvability of the linguistic system, e.g. via keeping a large number of linguistically equivalent pronunciation variants within the population. Thus, just as for biological systems, robustness and evolvability are intrinsically linked [14, 35, 62, 77].

Here, it is relevant to note that even though variation within the neutral spaces does not affect communicative outcomes right away, the variation is still perceived [68, 69] and can be copied by other language users (e.g. [79]). So, a slight change in pronunciation that leaves meaning unchanged could be adopted by other users of the same language for social reasons – e.g. group membership or status. If the variation within the neutral space created through quantality and categorical perception were completely imperceptible, certain sub-regions of the neutral space may not be preferred for social reasons. Moreover, the fact that variants within the neutral space can be perceived and copied means that the underlying system may change in a neutral fashion without impeding communication. This is exactly parallel to neutral drift in the biological context.

Implications for theory and practice

This paper has argued that many of the principles underlying robustness in biological systems are also relevant for the case of linguistic systems. As conceptual tools, the technical

notions of robustness, degeneracy and neutrality can be seen as unifying threads that run across many different sets of experimental data, including the presence of cue multiplicity, quantality, and categorical speech perception. This goes beyond standard accounts of communication such as information theory [80], which states that redundancy may counteract noise. As argued above, mere redundancy might not be enough. Instead, what biologists call "degeneracy" might be a more valid explanatory concept for the robustness and evolvability of spoken language.

Moreover, we might, indeed, never find true redundancy in spoken language, just as much as we do not find true redundancy in biological systems [12, 34]. In fact, it appears to be a logical necessity that speech cues are degenerate and not redundant: in order to be called a distinct cue, an acoustic signal needs to be *either* spectrally diverse *or* located at a different point in time – both of which deviate from pure "redundancy of parts" and show distribution across the spectral and temporal domain.

The present account goes beyond a mere analogy between language and biology: it has practical implications too. First, linguists may look toward biology for existing models of robustness to understand why speech communication functions in so many different situations, and why spoken languages change so frequently. For example, neutral network models can easily be interpreted within a linguistic context [35]. Furthermore, the present account leads to explicit predictions that can be tested: For example, contrasts that are found to be relatively more robust in perception experiments should also have a higher degree of cue degeneracy. In investigating these matters, linguists can operationally define and quantify robustness in specific local contexts. An example of how to work with robustness as a quantifiable concept is given in [81], the authors of which show that the confusability of certain sounds in noise [82] is correlated with the role that the sounds play within the respective communicative systems. That is, sounds that are more easily confused carry less functional load. This highlights one way in which the concept of robustness can be given a concrete quantitative characterization as "confusability in noise" (or lack thereof). Other researchers are beginning to explore quantifiable versions of notions such as "degeneracy" in the context of syntax [83].

The concepts discussed in this paper are also important for theoretical developments in language evolution research, because the present account highlights some of the *preconditions* for other theories of language evolution. Keller [63], for example, highlights that language change is influenced by speakers having multiple communicative desires, such as the desire to be understood and the desire to be expressive. From the perspective of the present paper, a prerequisite of satisfying these diverse communicative goals is a system that has multiple neutral solutions to the problem of "being understood", allowing these solutions to acquire different social values and serve different expressive functions. Similarly, Croft [7], among others, argues that much of the propagation of linguistic variants through a population of speakers happens because of the social values attached to these variants. This, too, necessitates multiple variants that can carry such social value without stopping the flow of communication. Thus, the present account in many ways outlines some of the conditions that need to be met in order for other theories of language evolution to work. Neutrality is crucial for understanding why language change is possible at all.

However, the present approach does not only benefit linguists. The close analogies between language and biology pointed out above show that language can in principle be used as a study system to understand general properties of robustness; properties that are, perhaps, independent of the physical make-up of the system. As "mutation rates" in language are much higher than in many biological study systems, and as language change has been successfully simulated in laboratory settings [84], some ideas involving robustness and evolvability are perhaps more easily and cheaply tested in language than in biology.

Such empirical investigations can, ultimately, also lead to a cross-systems theory of robustness. Many researchers have made explicit comparisons between robustness in biology and robustness in engineering [12, 14, 34], highlighting important similarities and differences between these two disparate domains. Spoken language is located somehow in between engineered and biologically evolved systems. It is, as argued by Keller [63], a "phenomenon of the third kind": language is within the domain of human intention because we use it, intentionally, to serve communicative functions; but language itself is – on the whole – not changed in an intentional fashion, but rather as a result of repeated use, learning and changing context. As a system that is between engineered and biologically evolved systems, language could perhaps be key to developing a system-wide theory of robustness.

Conclusions

The present paper points toward many principles of robustness being shared across seemingly dissimilar systems, including degeneracy, neutrality and evolvability. Both biological systems and linguistic systems are challenged by perturbations, and both systems meet these challenges in structurally analogous ways. This suggests a new point of synthesis between the fields of biology and linguistics. Even though biological evolution and linguistic evolution have distinct physical substrates, at an abstract systemic level, the commonalities between language and biology run deep.

The author has declared no conflict of interest.

References

1. **Ohala J**. 1989. Sound change is drawn from a pool of synchronic variation. In Breivik LE, Jahr EH, eds; *Language Change: Contributions to the Study of Its Causes*. Berlin: Mouton de Gruyter. p. 173–98.
2. **Johnson K, Mullennix JW**. 1997. *Talker Variability in Speech Processing*. San Diego: Academic Press.

Problems & Paradigms

3. **Hay J, Drager K**. 2007. Sociophonetics. *Annu Rev Anthropol* **36**: 89–103.
4. **Warren RM**. 1970. Perceptual restoration of missing speech sounds. *Science* **167**: 392–3.
5. **Schleicher A**. 1863. *Die Darwinische Theorie und die Sprachwissenschaft* [Darwin's Theory and Linguistics]. Leipzig: Hermann Böhlau.
6. **List JM, Nelson-Sathi S, Geisler H, Martin W**. 2013. Networks of lexical borrowing and lateral gene transfer in language and genome evolution. *BioEssays* **36**: 141–50.
7. **Croft W**. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow: Longman.
8. **Blevins J**. 2004. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
9. **Gontier N**. 2011. Depicting the tree of life: the philosophical and historical roots of evolutionary tree diagrams. *Evolution: Education and Outreach* **4**: 515–38.
10. **Gray RD, Jordan FM**. 2000. Language trees support the express-train sequence of Austronesian expansion. *Nature* **405**: 1052–5.
11. **Nelson-Sathi S, List JM, Geisler H, Fangerau H**, et al. 2011. Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proc Biol Sci* **278**: 1794–803.
12. **Kitano H**. 2004. Biological robustness. *Nature* **5**: 826–37.
13. **Carlson JM, Doyle J**. 2002. Complexity and robustness. *Proc Natl Acad Sci USA* **99**: 2538–45.
14. **Wagner A**. 2005. *Robustness and Evolvability in Living Systems*. Princeton: Princeton University Press.
15. **Conant GC, Wagner A**. 2004. Duplicate genes and robustness to transient gene knock-downs in *Caenorhabditis elegans*. *Proc Biol Sci* **271**: 89–96.
16. **Wagner A**. 2000. Robustness against mutations in genetic networks of yeast. *Nat Genet* **24**: 355–61.
17. **Kelso JS, Tuller B, Vatikiotis-Bateson E, Fowler CA**. 1984. Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *J Exp Psychol Hum Percept Perform* **10**: 812–32.
18. **Weismer G, Bunton K**. 1999. Influences of pellet markers on speech production behavior: acoustical and perceptual measures. *J Acoust Soc Am* **105**: 2882–94.
19. **Fowler CA, Turvey MT**. 1980. Immediate compensation in bite-block speech. *Phonetica* **37**: 306–26.
20. **Tremblay S, Shiller DM, Ostry DJ**. 2003. Somatosensory basis of speech production. *Nature* **423**: 866–9.
21. **Hughes OM, Abbs JH**. 1976. Labial-mandibular coordination in the production of speech: implications for the operation of motor equivalence. *Phonetica* **33**: 199–221.
22. **Lane HL, Tranel BW**. 1971. The Lombard sign and the role of hearing in speech. *J Speech Lang Hear Res* **14**: 677–709.
23. **Jones JA, Munhall KG**. 2000. Perceptual calibration of F0 production: evidence from feedback perturbation. *J Acoust Soc Am* **108**: 1246–51.
24. **Kasturi K, Loizou PC, Dorman M, Spahr T**. 2002. The intelligibility of speech with "holes" in the spectrum. *J Acoust Soc Am* **112**: 1102–11.
25. **Warren RM, Riener KR, Bashford JA, Brubaker BS**. 1995. Spectral redundancy: intelligibility of sentences heard through narrow spectral slits. *Percept Psychophys* **57**: 175–82.
26. **Strange W, Jenkins JJ, Johnson TL**. 1983. Dynamic specification of coarticulated vowels. *J Acoust Soc Am* **74**: 695–705.
27. **Félix MA, Wagner A**. 2006. Robustness and evolution: concepts, insights and challenges from a developmental model system. *Heredity* **100**: 132–40.
28. **Evans N, Levinson SC**. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Behav Brain Sci* **32**: 429–48.
29. **Calude AS, Pagel M**. 2011. How do we use language? Shared patterns in the frequency of word use across 17 world languages. *Philos Trans R Soc Lond B Biol Sci* **366**: 1101–7.
30. **Hamann S**. 2004. Retroflex fricatives in Slavic languages. *J Int Phonetic Assoc* **34**: 53–67.
31. **Pierrehumbert JB**. 2001. Exemplar dynamics: word frequency, lenition and contrast. In Bybee JL, Hopper PJ, eds; *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins. p. 137–58.
32. **Kingston J**. 2011. Tonogenesis. In van Oostendorp M, Ewen CJ, Hume EV, Rice K, eds; *The Blackwell Companion to Phonology*. Oxford: Wiley-Blackwell. p. 2304–33.
33. **Kingston J**. 2005. The phonetics of Athabaskan tonogenesis. In Hargus S, Rice K, eds; *Athabaskan Prosody*. Amsterdam: John Benjamins. p. 137–84.
34. **Edelman GM, Gally JA**. 2001. Degeneracy and complexity in biological systems. *Proc Natl Acad Sci USA* **98**: 13763–8.
35. **Whitacre J, Bender A**. 2010. Degeneracy: a design principle for achieving robustness and evolvability. *J Theor Biol* **263**: 143–53.
36. **Mason P**. 2010. Degeneracy at multiple levels of complexity. *Biol Theory* **5**: 277–88.
37. **Price CJ, Friston KJ**. 2002. Degeneracy and cognitive anatomy. *Trends Cogn Sci* **6**: 416–21.
38. **Noppeney U, Friston KJ, Price CJ**. 2004. Degenerate neuronal systems sustaining cognitive functions. *J Anat* **205**: 433–42.
39. **Stevens KN, Keyser SJ**. 2010. Quantal theory, enhancement and overlap. *J Phonetics* **38**: 10–9.
40. **Hawkins S**. 2010. Phonological features, auditory objects, and illusions. *J Phonetics* **38**: 60–89.
41. **Kingston J, Diehl RL**. 1994. Phonetic knowledge. *Language* **70**: 419–54.
42. **Wright R**. 2004. A review of perceptual cues and robustness. In Steriade D, Kirchner R, Hayes B, eds; *Phonetically Based Phonology*. Cambridge: Cambridge University Press. p. 34–57.
43. **Lisker L, Abramson AS**. 1964. A cross-language study of voicing in initial stops: acoustical measurements. *Word* **20**: 384–422.
44. **Moreton E**. 2004. Realization of the English postvocalic [voice] contrast in F1 and F2. *J Phonetics* **32**: 1–33.
45. **Benkı̓ JR**. 2001. Place of articulation and first formant transition pattern both affect perception of voicing in English. *J Phonetics* **29**: 1–22.
46. **Haggard M, Ambler S, Callow M**. 1970. Pitch as a voicing cue. *J Acoust Soc Am* **47**: 613–7.
47. **Raphael LJ**. 1972. Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *J Acoust Soc Am* **51**: 1296–303.
48. **Lisker L**. 1957. Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* **33**: 42–9.
49. **Repp BH**. 1979. Relative amplitude of aspiration noise as a voicing cue for syllable-initial stop consonants. *Lang Speech* **22**: 173–89.
50. **Hawkins S, Nguyen N**. 2004. Influence of syllable-coda voicing on the acoustic properties of syllable onset /l/ in English. *J Phonetics* **32**: 199–231.
51. **Parker EM, Diehl RL, Kluender KR**. 1986. Trading relations in speech and nonspeech. *Percept Psychophys* **39**: 129–42.
52. **Wagner A**. 2005. Distributed robustness versus redundancy as causes of mutational robustness. *BioEssays* **27**: 176–88.
53. **Xu Y, Xu CX, Sun X**. 2004. On the temporal domain of focus. *Proc 2nd Int Conf Spreech Prosody* 81–4.
54. **Jusczyk PW, Cutler A, Redanz NJ**. 1993. Infants' preference for the predominant stress patterns of English words. *Child Dev* **64**: 675–87.
55. **Fernandes T, Ventura P, Kolinsky R**. 2007. Statistical information and coarticulation as cues to word boundaries: a matter of signal quality. *Percept Psychophys* **69**: 856–64.
56. **Saffran JR, Newport EL, Aslin RN**. 1996. Word segmentation: the role of distributional cues. *J Memory Lang* **35**: 606–21.
57. **Shatzman KB, McQueen JM**. 2006. Segment duration as a cue to word boundaries in spoken-word recognition. *Percept Psychophys* **68**: 1–16.
58. **Ettlinger M, Finn AS, Hudson Kam CL**. 2011. The effect of sonority on word segmentation: evidence for the use of a phonological universal. *Cogn Sci* **36**: 655–73.
59. **Christiansen MH, Allen J, Seidenberg MS**. 1998. Learning to segment speech using multiple cues: a connectionist model. *Lang Cognitive Proc* **13**: 221–68.
60. **Silva DJ**. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* **23**: 287–308.
61. **Kirby J**. 2013. The role of probabilistic enhancement in phonologization. In Yu A, ed; *Origins of Sound Change: Approaches to Phonologization*. Oxford: Oxford University Press. p. 228–46.
62. **Wagner A**. 2005. Robustness, evolvability, and neutrality. *FEBS Lett* **579**: 1772–8.
63. **Keller R**. 1990. *Sprachwandel: Von der unsichtbaren Hand in der Sprache* [On language change: The invisible hand in language]. Tübingen: Francke Verlag.
64. **Stevens KN**. 1989. On the quantal nature of speech. *J Phonetics* **17**: 3–46.
65. **Diehl RL**. 2008. Acoustic and auditory phonetics: the adaptive design of speech sound systems. *Philos Trans R Soc Lond B Biol Sci* **363**: 965–78.
66. **Chi X, Sonderegger M**. 2007. Subglottal coupling and its influence on vowel formants. *J Acoust Soc Am* **122**: 1735–45.
67. **Liberman AM, Harris KS, Hoffman HS, Griffith BC**. 1957. The discrimination of speech sounds within and across phoneme boundaries. *J Exp Psychol* **54**: 358–68.

68. **McMurray B, Aslin RN, Tanenhaus MK, Spivey MJ**, et al. 2008. Gradient sensitivity to within-category variation in words and syllables. *J Exp Psychol Hum Percept Perform* **34**: 1609–31.

69. **McMurray B, Tanenhaus MK, Aslin RN, Spivey MJ**. 2003. Probabilistic constraint satisfaction at the lexical/phonetic interface: evidence for gradient effects of within-category VOT on lexical access. *J Psycholinguist Res* **32**: 77–97.

70. **Daniels BC, Chen YJ, Sethna JP, Gutenkunst RN**, et al. 2008. Sloppiness, robustness, and evolvability in systems biology. *Curr Opin Biotechnol* **19**: 389–95.

71. **Masel J, Siegal ML**. 2009. Robustness: mechanisms and consequences. *Trends Genet* **25**: 395–403.

72. **Cutting JE**. 1982. Plucks and bows are categorically perceived, sometimes. *Percept Psychophys* **31**: 462–76.

73. **Kuhl PK, Miller JD**. 1975. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. *Science* **190**: 69–72.

74. **Dooling RJ, Okanoya K, Brown SD**. 1989. Speech perception by budgerigars (Melopsittacus undulatus): the voiced-voiceless distinction. *Percept Psychophys* **46**: 65–71.

75. **Wyttenbach RA, May ML, Hoy RR**. 1996. Categorical perception of sound frequency by crickets. *Science* **273**: 1542–4.

76. **Eimas PD, Siqueland ER, Jusczyk P, Vigorito J**. 1971. Speech perception in infants. *Science* **171**: 303–6.

77. **Masel J, Trotter MV**. 2010. Robustness and evolvability. *Trends Genet* **26**: 406–14.

78. **Wedel A**. 2006. Exemplar models, evolution and language change. *Linguist Rev* **23**: 247–74.

79. **Goldinger SD**. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychol Rev* **105**: 251–79.

80. **Shannon CE**. 1948. A mathematical theory of communication. *Bell System Tech J* **27**: 379–423 & 623–56.

81. **Graff P**. 2012. Communicative Efficiency in the Lexicon. PhD thesis, Massachusetts Institute of Technology.

82. **Miller GA, Nicely PE**. 1955. An analysis of perceptual confusions among some English consonants. *J Acoust Soc Am* **27**: 338–52.

83. **Van de Velde F**. 2014. Degeneracy: the maintenance of constructional networks. In Boogart R, Colleman T, Rutten G, eds; *Constructions All the Way Everywhere: Extending the Scope of Construction Grammar*. Berlin: Mouton de Gruyter.

84. **Scott-Phillips TC, Kirby S**. 2010. Language evolution in the laboratory. *Trends Cogn Sci* **14**: 411–7.

Problems & Paradigms