Phylogenetic inference for function-valued traits: speech sound evolution

The Functional Phylogenies Group

Corresponding Author: Nick Jones nicholas.jones1@imperial.ac.uk

John A.D. Aston - Warwick University, Statistics, UK
Dorothy Buck - Imperial College London, Mathematics, UK
John Coleman - Oxford University, Phonetics Laboratory, UK
Colin J. Cotter - Imperial College London, Aeronautics, UK
Nick S. Jones – Imperial College London, Mathematics, UK
Vincent Macaulay - Glasgow University, Mathematics and Statistics, UK
Norman MacLeod – Natural History Museum, Palaeontology, UK
John M. Moriarty - Manchester University, Mathematics, UK
Andrew Nevins - University College London, Linguistics]

Abstract: Phylogenetic models have recently been proposed for data that are best represented as a mathematical function (function-valued). Such methods can used to model the change over time in function-based descriptions of various data of interest to evolutionary biologists, including the sound of speech. This approach to phylogenetic inference and analysis is challenging, both in terms of modelling the phylogenetics of functions, and in engaging with previously existing evidence for character-state change. Nevertheless, it is both a real and exciting prospect. Our approach could provide those interested in investigating a greater range of
evolutionary processes with the ability to utilize statistical hypothesis testing procedures and to create estimates of the states of function-valued characteristics (e.g., speech sounds) at earlier historical times.

**Beyond symbols**

The word 'functional' has multiple meanings. One of these — and the one we will focus on in this article — is 'represented by a continuous mathematical function'.

Given a set of traits that can be described as functions (these are called function-valued traits [1]) and that have been mapped to positions on a network of phylogenetic relationships, what can be said about values of these traits at other points on the phylogeny? While the study of the evolution of both genetic sequences and quantitative characters, in particular comparative method analysis, has seen extensive application and technical development, the study of the phylogenetics of function-valued characteristics is comparatively less advanced.

Herein we propose methods that could be useful in extending the ability of evolutionary inferences in many areas of evolutionary system studies. Function-valued traits are quite common and appear across many areas of biology, from human height vs age to backfat thickness of animals vs weight. However, the idea that traits can be represented as functions is new to other fields less closely associated with biodiversity studies. One such field is linguistics. Accordingly, speech sound evolution will serve as a running example of our proposed approach.

As background to this choice we note that this is a time of great promise in linguistic cladistics. For example, phylogenetic methods have been used to shed light on the early differentiation of Indo-European languages and the dynamics of the
appearance of new languages [2,3]. There have also been agent-based simulations
investigating how speech sounds might evolve, e.g. [4].

One method for dealing with function-valued traits in a phylogenetic context is the
functional phylogenetic regression method recently proposed by Jones and Moriarty
[5]. This data analysis technique differs from previous approaches in its use of
continuous characters [6,7,8] and by modelling evolved traits explicitly as functions.
Building upon recent work in machine learning [9], this mode of phylogenetic
regression takes account of constraints on the patterns of variation for function-valued
traits over a phylogeny that are not captured by more traditional multivariate
models. Since this approach avoids summarizing these data it also supports the
prediction of ancestral function forms directly rather than through the estimation of a
few of their characters. For speech-sound analysis we could model each articulatory
movement as a sequence of positions of the vocal apparatus, a continuous view of
which was achieved in a straightforward manner. In continuing our efforts to develop
this approach our object for enquiry is simply the speech sound itself (or functional
versions of it such as the spectrogram or empirical covariance surface [see
Glossary]) which can be treated as a mathematical function.

Our primary goal is to discuss how quantitative methods can allow us to infer
probability distributions over possible ancestral function-valued traits (see Fig. 3).
Samples from these distributions offer the prospect of estimating functional
descriptions of speech from languages which have no living speakers. Provided this
prospect can be realized rigorously, such an ability would be a substantial advance
from symbolic phylogenetic modelling towards (in this case) true acoustic
reconstruction, yielding a new tool in comparative philology. In a larger sense, however, this function-based approach has relevance for many types of biological data and, indeed, for phylogeny reconstruction itself.

The approaches proposed below lend themselves to testing hypotheses about the sounds of past languages suggested, for example, by textual inference or from the written descriptions of phoneticians in the ancient world e.g. [10,11,12,13]. Our attempts to realize this aim are of course quite preliminary. Nevertheless, we are intrigued by the prospect of using phylogenetic methods to enable us to hear approximations to the sound of dead languages. This goal is most naturally approached through the analysis of mathematical functions describing actual speech. Fortunately, we are close to having all the materials and mathematical methods needed to make the first attempts toward this goal’s realization.

**Speech sounds as function-valued traits**

Speech sounds have a wide range of component frequencies, from a few hertz (e.g., speaking vs. silence) to c. 6 kHz, with a dynamic range of c. 60 dB. Despite this variability, speech sound waves are fairly similar over time intervals of 5-10 ms or more, the higher frequencies arising from multiples of the fundamental frequency (see Glossary) in voiced sounds, and resonances of the vocal tract. Therefore, it is possible - indeed, now routine in many areas of speech processing - to decompose speech sound waves into time series of multiple (c. 12-20), slowly-varying components which are approximately constant on a timescale of 5-10ms. As a result, the representation of speech sound as a function, a function-valued trait, or set of
functions is common within phonetics and speech technology e.g. in Linear Prediction Coding [14].

Words as functions

Speech sounds have far-reaching contextual effects on one another (= co-articulation [15]). It is therefore very difficult to decompose recordings of speech into separate, consecutive ‘sounds’ with a high degree of objectivity and repeatability. Although symbolic representations such as phonetic transcriptions are based on separate symbols, in the continuum of articulated speech, sounds merge into one another [16,17]. As we will discuss below, historical and evolutionary sound changes rarely affect individual component sounds. Rather, sounds evolve together within the context of specific words. Accordingly, we may sidestep the practical difficulties of segmenting words into separate sounds by modelling the evolution of (sets of) whole words. The challenge of segmenting vocalizations becomes even more pronounced when considering animal sounds e.g. [19,20,21]. A functional approach to the analysis of these data is not only the most appropriate for the analysis of human speech sounds, it is also the most generalisable.

Can speech sound change can be regarded as incremental?

Language change is the result of imperfect learning, reanalysis, gradual changes in pronunciation, and/or recurrent misperception of particular sequences [see e.g. 22]. Consider the differences between Spanish, Catalan and Portuguese, three languages that originate from a common ancestor (Iberian, itself descended from
Latin). For example, the *bl*- sound is observed in two of the three languages (Sp. *blanco*, Cat. *blanc*) and *br*- in one of them (Port. *branco*), as well as *bl*- being in the original ancestor of all three. This change, from *l* -> *r*, is necessarily context-sensitive. It did not occur across-the-board in these languages — in fact, it never occurs at the beginning of a word (cf. Sp. *libro*, Cat. *llibre*, Port. *livro*), but only after a preceding consonant in the same syllable. This change pertains specifically to the similarity of *bla*- to *bra*- in two crucial regards: first, it shows that the two are similar enough that such a change is gradual and minimal in terms of the articulators involved and second, in the perceived acoustic similarity, rather than a wholesale change of, say, *bla*- to *shroo*-.. The comparison between words beginning with *bl*- vs. *br*- in different branches of the Iberian languages represents the result of accumulated and recurrent sound changes reflecting the long-term dynamics of repeated listener/speaker mislearning, misperception, mispronunciation and cognitive reorganization.

Rather than occurring at the level of single consonants and vowels (e.g., *l* changes to *r*), could it be the case that continuous changes to the entire word are a better way of understanding speech sound change? Consider one of the most common linguistic changes, namely substitution. While it is fruitful to formalize the patterns of change in a language as operations on strings, symbolic rules such as *xyz* → *xzz* do not always capture the fact that sound changes are non-arbitrary: sound substitutes are similar sounds. For example, Latin *clavis* (with initial *[k]*-) became Italian *chiave* *[kjave]*; *civitas*, (also with initial *[k]*-), became Italian *citta* (*[tʃitta]*) and French *cité*, with initial *[s]*. Numerous examples of this chain of sound changes (in the phonetic alphabet *k* → *kj* → *cç* → *tʃ* → *ts* → *s*) can be found across many language families.
A functional perspective on sound change

To understand such historical processes, it is necessary to make the notion of ‘similarity’ more precise. Two articulatory reasons for substitutions are undershoot and overshoot (see Glossary and Fig. 1A and 1B).

Figure 1A

![Figure 1A](image)

Figure 1B

![Figure 1B](image)
Figure 1A. Here the y-axis is the size of the aperture of the oral airway, i.e. of the vocal tract.

**Undershooting a closing movement:** e.g. \([t] \rightarrow [s]\), the ‘t’ in *permit* becomes the ‘s’ in *permission* (noncontinuant sounds (stops) become continuants) or Latin *spatium* becomes French *espace* and English *space*.

Blue line: An articulatory movement that completely closes the vocal tract (aperture = 0 from \(t = 100\) to \(t = 250\)), as for \([t]\). Red line: An incomplete closing movement, as for \([s]\). The vocal tract is constricted, but before complete closure is reached, the tract re-opens.

**Undershooting an opening movement:** e.g. a low (open) vowel becomes less wide open, as in *telegr[a]ph* \(\rightarrow\) *telegr[a]phy* \([a] \rightarrow [ə]\). Blue:

An opening-closing movement of the vocal tract, as for \([a]\). Red: An incompletely executed opening movement, as for \([ə]\). Perhaps under a time constraint, the vocal tract begins to close again at \(t = 150\), before it has reached the target aperture.

The two other types of sound change, deletion and insertion, can also be understood in terms of functions. As an operation on strings, deletions are of the form \(\ldots axb\ldots \rightarrow \ldots ab\ldots\) but, viewed physically and continuously, we can consider deletion to be the gradual shortening of \(x\) until its duration becomes negligible. For example,

progressive shortening of vowels to the point at which they have no duration gives reduced pronunciations (e.g., suppose \(\rightarrow\) s’pose). The ubiquitous redundancy in speech allows deletion to occur without significant loss of information. Viewed as an operation on strings, the addition or insertion of sounds has them appear from nowhere: \(\ldots ab\ldots \rightarrow \ldots ayb\ldots\) From a functional perspective inserted sounds are often comprehensible in terms of their context and can be sometimes associated with a misphasing phenomenon.
We have thus seen that we can reinterpret attested sound changes – deletions, insertions, and substitutions – in terms of continuous changes to continuous functions, rather than discrete and arbitrary transformations of strings.

**Analysis of ‘functionalized’ speech sounds**

Having argued that speech sounds and their evolution can be treated in terms of mathematical functions, we are now in a position to describe the tools that enable us to use mathematical functions as characters in a phylogenetic analysis.

**Box 1 Gaussian Process Regression**

Examples of Gaussian processes are the Brownian motion and the Ornstein-Uhlenbeck processes [9] (see Fig. 2A). The latter can be viewed, informally, as a Brownian motion which tends towards a fixed mean. Both Brownian motion and Ornstein-Uhlenbeck processes have been used to describe and analyse continuous characters as part of the comparative method in phylogenetics [6-8]; they are used as models of how univariate continuous traits might evolve in time. It is worth noting that, while such continuous traits are normally used in approaches that assume a fixed phylogeny (e.g., the comparative method), they can, in principle, be used for phylogenetic inference. In this contribution we are interested primarily in how function-valued data (not only univariate data) evolve in time.

Linear regression fits a function \( a \cdot l + b \) (where \( a \) and \( b \) are fitting parameters) to \( N \) data values \( \{y(l_1), y(l_2), y(l_3) \ldots y(l_N)\} \) measured at some \( N \) co-ordinates of observation\( \{l_1, l_2, l_3 \ldots l_N\} \). Gaussian process models can also be used to fit function-
valued data. In a peculiar fashion, Gaussian process regression presumes a fit of a multivariate normal or Gaussian distribution to data: a multivariate Gaussian where each co-ordinate of observation in a set of $N$ such measurements, \{\(l_1, l_2, l_3 \ldots l_N\)\}, is assigned a variable in the $N$-dimensional distribution. The data values, \(y\), at these points, \(\{y(l_1), y(l_2), y(l_3) \ldots y(l_N)\}\) are just one sample from this distribution. Just like any multivariate Gaussian these processes are specified by how the variables co-vary (the matrix of covariances): different types of covariances, or couplings, between the co-ordinates of observation yield different process models. In Figure 2A the points on the blue curve covary differently from points on the black curve. Gaussian process models are called non-parametric because, unlike parameterized models (like \(a.l + b\)) the number of parameters increases with the number of observations \(N\). A final relevant property of Gaussian processes is that they can be used, given knowledge of data values at some co-ordinates of observation \(\{y(l_1), y(l_2), y(l_3) \ldots y(l_N)\}\), to make predictions about values at other co-ordinates \(\{r_1, r_2, r_3 \ldots \}\) about which we do not have any data (see Fig. 2b). Box II considers a generalisation of this prediction task to unmeasured data at different evolutionary times.

Figure 2. (A) Example Gaussian Process samples - Black line: a sample from a Brownian Motion process. Red line: a sample from an Ornstein-Uhlenbeck process. Blue line: a sample from a process with a squared-exponential covariance function.
Processes like these can be fit to data easily. (B) Making inferences about unseen data within a function-valued trait: A missing data problem. The data are represented by the black curve (we assume the data were generated by a Gaussian process). The red curve represents the locus of expected values at the missing data points and the blue curves are the one standard deviation uncertainties at each value. The black-dashed curve is a sample from the process during this period of missing data. In Box 2 we will treat unseen ancestral functions as missing data.

The treatment of data as discretely observed mathematical functions or curves has received substantial attention in the recent mathematical, statistical and machine learning literature. In Functional Data Analysis [23,24] the data are often assumed to arise from a stochastic process which yields smooth functions (possibly subject to noise). Again, note that the word 'functional' here, and in the following, refers to mathematical functions, not an aspect of biological function. Considerable progress has been made in extending standard statistical notions that are in common use for univariate or multivariate data (e.g., regression, classification) to these functional data objects [25]. Box 1 outlines one type of functional data object: a Gaussian process. Further, recent work in quantitative genetics has applied functional data analysis to function-valued traits [26,27]. However, little work has been done to date using functional data in phylogenetic inference, particularly in a non-parametric context.

Explicitly functional techniques have been applied in linguistic analysis. Some of the earliest research on functional data analysis corresponded to tracking articulatory lip movement trajectories during speech [28]. A recent empirical study has shown that,
in carefully controlled conditions, Functional Data Analysis can help classify different yet closely related consonant sounds [29]. Some researchers are now taking a functional approach to word contractions like the suppose→s’ppose example discussed above e.g. [30]. Several studies have considered functional representations of $f_0$ curves ($f_0$, the fundamental frequency, is the acoustic basis of perceived voice pitch; see Glossary). Grabe et al. [31] used a component set of functions to examine $f_0$ intonation contours. Changes in intonation (tones) which can impart meaning to words are present in many of the worlds’ languages, including languages such as Mandarin Chinese. Non-parametric functional data analysis has been used [32] to characterise the tones of Qiang, a Sino-Tibetan language. It was shown, first, that known tonal $f_0$ patterns can be reproduced from the empirical covariance surface (see Glossary) and, second, that additional patterns, not necessarily postulated from isolated utterances can be identified as relevant. This was then found to apply to a much larger body of $f_0$ curves in Mandarin [33].

**Evolutionary Inference for function-valued data**

The phylogenetic Gaussian process models that we have outlined in boxes 1 and 2 describe one way in which we could perform evolutionary inference with function-valued data. Figure 3 provides an example of attempting to predict an ancestral value (which in the case of speech, can be tested against our other knowledge of past speakers, see below). But we note that one can also use this approach to support direct phylogenetic inference. Gaussian processes make a minimal set of assumptions about the patterns which they model. These assumptions are that the distribution of each value of each trait is Gaussian, and all correlations between time-
points in the signal are specified only by how pairs of points covary (the data is assumed to have only pairwise correlations). Though these might seem to be strong constraints, the Gaussian process framework is remarkably flexible and robust when applied to a wide variety of inference tasks [9].

**Box 2 Gaussian Processes for Evolutionary Inference**

In Box 1 we noted that a Gaussian process’s nature (just like any multi-variate Gaussian) is specified by how the co-ordinates of observation, \(\{l_1, l_2, l_3, \ldots l_N\}\), co-vary. We have given examples where the set of co-ordinates of observations corresponds to observations or measurements at successive points in time. In fact, these points of observation need not only be in one dimension: they could be observation co-ordinates of a 2D function on a plane (e.g., the \(x-y\) co-ordinates on a surveyor’s topographical map).

Let us suppose we call one dimension evolutionary time and the other dimension trait-space. A 2D function can then be viewed as a 1D function with spatial co-ordinates (a function-valued trait) evolving in time. If we think of the curve changing its form down an evolutionary tree (see Figure 3), any path from tree root to branch tip can be viewed as a 2D function. Jones and Moriarty [5] take this intuition and work out the correct form for the covariances between points of observation anywhere on any given tree \(T\) provided we know — or can assume — how co-ordinates of observation on the path from root to tip covary. The degree of covariance depends on the phylogeny. Different phylogenies, \(T_1\) and \(T_2\), will yield different co-variances because the relative positions of co-ordinates of observation, \(\{l_1, l_2, l_3, \ldots l_N\}\), will
change. Given knowledge of the covariances one can formulate likelihoods for the different tree topologies, calculate appropriate Bayes factors, and so determine which phylogeny, $T_1$ or $T_2$, is most consistent with the observed data.

Once we understand how traits at different points on the tree covary, we can also develop a distribution for trait values at internal points and/or nodes on the tree Fig. 3. For example, using such a distribution it would be possible to attempt to forecast trait values at any points in the tree with co-ordinates $\{r_1, r_2, r_3\ldots\}$, given only observations at the tree’s tips $\{y(l_1), y(l_2), y(l_3)\ldots y(l_N)\}$ in exactly the same way as we approached the missing data problem in Fig 2B. It is also possible to obtain a sample from a distribution for trait values at internal points to yield example ancestral functions.
Fig. 3. Making inferences about unseen data pertaining to an ancestral function-valued trait: Phylogenetic Gaussian process regression. The black lines are the function-valued data we observe at the leaves of a known evolutionary tree and we assume a notional curve in the deep past which is flat. Two curves have a common ancestor at two units of evolutionary time and all three have a common ancestor at three units of evolutionary time. We construct a probability distribution over function-valued traits at 5 units of evolutionary time (we do this by modelling the data as having a covariance in the space of the function-valued trait like that of the smooth blue curve in Fig. 2a and having an Ornstein-Uhlenbeck covariance in evolutionary time, like the red curve in 2a and we also allow a small amount of noise). The red line is the expected value of the ancestral function at each point in the trait space and the blue lines are one standard deviation uncertainties in that value. The black dots are a sample of the Gaussian process at the earlier time, conditional on the data at the tips. Note that we could alternatively have attempted to infer the phylogeny most consistent with the function-valued tip data. We have placed a notional example next to this for a speech sound evolution.

The use of the full Gaussian process framework described in the boxes is just one route to evolutionary inference with function-valued data. Another option would be to pool each piece of functional data available and use these together to construct a small set of component functions that are appropriate for representing any member of the set (e.g., obtaining a set of basis functions by using a functional version of principal components analysis [23,24]). This would yield a natural set of functions with which to decompose any one of the function-valued traits; i.e., any function can then be viewed as a weighted combination of these component functions. As a trait varies through evolutionary time it could be that this set of component functions is unchanging. In this way change might be described as occurring only in the relative weighting of each component function (its coefficient in the functional basis we have extracted from the data) needed to represent the trait. Since now we are only studying the pattern of changes in a vector of weights, methods that model the
evolution of univariate and multivariate characters [6-9] can be used rather than the full Gaussian process framework described in the boxes. We note that these approaches still require a choice of covariance structure between points in time. Again these methods can help infer phylogenies and/or generate hypotheses about ancestral function-valued traits.

**Evolutionary inference for speech-sound data**

With the aim of performing function-valued evolutionary inference with speech-sound data the best functional representation of these data remains open: we suggest three possibilities. Considering intonation contours as evolving function-valued traits might shed light on the relationships between tonal and non-tonal versions of speech sounds. Spectrograms, which show how the different frequency components present in the signal vary through time, are useful (2D) functional representations of speech sounds since these allow approximate reconstruction of the speech sound from their forms; a similar case holds for coefficients derived from Linear Prediction Coding as mentioned above.

Possible inhomogeneities in speech-sound evolution include abrupt changes, and punctuated bursts of evolution [2]: we note that approaches to these possibilities have already been considered in the context of single-character and multivariate phylogenetic models of evolution [35]. Such approaches correspond, in our context, to allowing the covariances between points in time to themselves be time-varying. Although we regard these proposed approaches as crude (as first models should be), a standard approach to compensate for a simplistic model is to consider traits
which are clearly homologous as this allows one to partly control for more exotic forms of evolution not (presently) included in the model.

The evolution of speech may be considered at several different levels of resolution: phonemes, words, or word groups. Historical linguists often consider the co-evolution of groups of words which have phonetic similarities. In a Gaussian process framework, co-evolution could be modelled by specifying appropriate covariances between the words. By incorporating more information such joint approaches offer the prospect of facilitating explicit hypothesis tests concerning the character of individual ancestral spoken words.

Directions and Data

Now that we have made the case for a function-based treatment of data such as speech and sound change, and noted ways in which existing methods and possible variants could be used to make data-driven models for spoken language evolution, it is appropriate to discuss more general questions of speech sound evolution data and method. There currently exist several relatively uncontroversial linguistic phylogenies, a natural starting point being the Indo-European languages [37,38,39]. These can be used for a first approximation to the evolutionary history of individual spoken words. Contemporary samples of speech from languages at the most recent branches of the linguistic evolutionary tree are readily available, e.g. [40].

Though it is not required for our approach, recordings of experts in ancient languages - even though they are not native speakers, of course - could yield root
speech recordings useful for testing proximate hypotheses of speech-sound evolution. Similar (homologous) words across languages have already been found for a large number of words, by centuries of philological scholarship (e.g., [39]). Existing methods can model (and so accommodate) variation in the duration of recorded speech sounds through the use of warping and other registration algorithms (e.g. [41,23]). From the tree and from tip-word recordings, it could thus be feasible to infer and hence synthesize the sounds of words from any desired generation, along any branch, through a statistical reconstruction of traditional philological methods.

When analyzing empirical data, it is possible first to pool a training dataset to identify the nature of variation. Random factors (e.g., in linguistics, different speakers), can be taken into account at this stage using methods akin to those commonly used in function valued trait analysis [34]. where explicit variational types (e.g., genetic and environmental variability in function valued trait models) are modelled separately, then combined. Variability may also be accounted for via mixed effects models, an extension of linear models already considered for function-valued data in linguistics [32]. Traits would then be subject to an evolutionary process that has been explicitly adjusted to the patterns already observed in empirical data.

We have discussed how function-valued data might be used to perform both phylogenetic inference and also to generate hypotheses about ancestral linguistic traits. We might conclude by answering the question: why would an ability to reconstruct past speech sounds be worthwhile? We believe that, as well as connecting us with our past, moderate success in this endeavour could have a substantial effect on historical linguistics. More broadly though, it is our contention
that these data-analysis approaches can make a positive contribution to the investigation of evolutionary processes in a large number of biological contexts. Whereas the treatment of continuous variables in the context of a phylogenetic analysis was once considered unacceptable [42, 43, 44], there is a new interest not only in using phylogenies to understand and test hypotheses involving the historical development of continuous traits (e.g., via the comparative method [8, 45, 46]), but also to allow continuous variables to participate directly in the inference of phylogenetic relationships among species. [47, 48, 49]. We see our efforts to create statistical tools that can be used to incorporate function-valued data into phylogenetic data-analysis contexts as closely related to this (now) well-established research programme. These methods hold the promise of extending the power of phylogenetic analysis into new research fields, bringing new data to bear on the many phylogenetic problems that have thus far resisted resolution under the phylogenetic systematics paradigm, and providing new ways of treating old data to maximize their contribution to the understanding of nature and of natural processes, including those traits that have a critical role in shaping the evolutionary history of our own species.

7. References


319, 588


427 http://catalog.elra.info
431 201-209.
435 Homology: The Hierarchical Basis of Comparative Biology (Hall, B.K., ed), pp. 197–
436 227, Academic Press
438 biology, Oxford University Press.
440 geometric interpretations. Evolution 55 (11), 2143–2160.
442 shape and phylogeny (MacLeod, N. and Forey, P.L., eds), pp. 100–138, Taylor &
443 Francis.
447 algorithms for landmark optimization Cladistics 27 (1), 42–51
448
449
450 **Glossary:**
\$f_0\$ - **Fundamental frequency.** From an acoustic (sound) perspective, a speech signal is a complex periodic wave composed of multiple sine waves. In voiced sounds, the lowest frequency of this complex wave is its \$f_0\$, the frequency at which the vocal folds are opening and closing.

**[Empirical] covariance function (empirical covariance surface)** - The covariance surface of a function is the functional analogue of a covariance matrix for multivariate data. It must be positive or at least non-negative definite and in general is assumed to have finite summable eigenvalues giving constraints on its nature (which must be preserved in any evolutionary process). The empirical covariance surface is the covariance surface estimated directly from multiple realisations of the underlying stochastic process (the data).

**Undershoot:** In a movement (or a dynamical model of that movement), "undershoot" means that the movement failed to meet the intended position target. For example, one may make attempt to grasp an object (e.g. a fruit hanging in a tree), but fail to grasp it. One may attempt to kick a moving ball, but be too slow to reach it. Or, in speaking, one may intend to open the mouth for a certain vowel, but because of the rapidity of speech fail to achieve the intended articulation.

**Overshoot:** In a movement (or a dynamical model of that movement), "overshoot" means that the magnitude of movement was excessive, given the intended position target. For example, one may reach for a cup, but accidentally knock it over instead. Or an infant, intending to say "sock", may through inexperience utter "dock" instead,
because they have overestimated how much to move their tongue tip.

(Bayesian) Nonparametric model: an approach to (Bayesian) model selection and adaptation in which the model size is allowed to grow with data size. In these techniques, while individual variables are typically assumed to belong to parametric distributions, the number of parameters is not fixed.

Gaussian process (GP): a stochastic process whose marginal distribution at each point is Gaussian.