2025. Proc Ling Soc Amer 10(1). 5934. https://doi.org/10.3765/plsa.v10i1.5934.

Dynamic Field Theory unifies discrete and continuous aspects of linguistic representations

Michael C. Stern[*]

Abstract. In recent years, a growing body of research has sought to explain linguistic phenomena in terms of the dynamics of neural activity, through the lens of Dynamic Field Theory (DFT: Schöner, Spencer & DFT Research Group 2016). DFT is a general framework for understanding perception, action, and cognition as resulting from activity in interconnected populations of neurons. DFT formalizes neural activity in the language of *nonlinear dynamical systems.* This expression allows apparently categorical behavior to emerge from an underlyingly continuous state space. In this paper, I provide a review of research investigating linguistic phenomena through the lens of DFT, with a particular emphasis on how this research unifies discrete and continuous aspects of linguistic representations, and in doing so, unifies disparate empirical findings and theoretical insights from various domains of linguistics.

Keywords. Dynamic Field Theory; dynamical systems; neurolinguistics; psycholinguistics; computational models

1. Discrete and continuous representations in linguistics. Discrete categories play an important role in describing language. Discrete categorical representations like distinctive features, phonemes, morphemes, and syntactic categories, in combination with operations like rules and constraints, have afforded explanations for a wide range of phenomena in human language. Theories that utilize discrete representations often make strong predictions about cross-linguistic typology, and place constraints on possible human languages, a core goal of linguistic theory. Moreover, computational models of language processing that make use of the discrete representations from linguistic theory have offered explanations of patterns in speech errors and speech planning times (e.g., Dell 1986; Levelt, Roelofs & Meyer 1999). In general, more similar representations are more likely to interact during speaking. For instance, an error of *substitution*—accidentally replacing one representation, e.g., phoneme, with another—is more likely when the intended and errorful representations are more similar. For example, if one intends to say <god>, they are more likely to accidentally say <cod> than they are to say <sod>, because /g/ and /k/ differ only in a single phonological feature, voicing, while /g/ and /s/ differ in a greater number of features. Computational models based on discrete representations encode similarity through hierarchical levels of representation. /g/ and /k/ are categorically distinct on a phonemic level of representation, but they share many connections to a featural level of representation. It is this pattern of overlap in connections that encodes similarity between discrete representations.

1.1 RELATIONSHIP BETWEEN DISCRETE REPRESENTATIONS AND CONTINUOUS MEASUREMENTS. Discrete linguistic representations, while theoretically powerful, are not directly observable. Speech articulation involves the movement of effectors like the tongue, lips, and larynx through continuous space over continuous time. These movements perturb the air around the speaker, generating pressure fluctuations (sound waves) which are also continuous in space and time. Moreover, the

---

[*] I would like to thank the members of the Yale Dynamics Group for useful feedback and discussion, especially Jason Shaw, Manasvi Chaturvedi, Maria Teresa Borneo, Alessandra Pintado-Urbanc, Ayla Karakaş, and Clara Terlaak. All errors are my own. Author: Michael C. Stern, Yale University (michael.stern@yale.edu).

neural activity underlying speech production and comprehension is continuous in state space: the membrane potentials of cells vary continuously over time, contributing to continuous patterns of aggregate activity across neural populations. Finally, the neural connectivity which underlies linguistic knowledge varies continuously, with the strength of synaptic connections between neurons changing gradiently over an individual's lifetime.

A logical possibility is that these continuous variables *instantiate* discrete linguistic representations, but ultimately only the discrete representations are relevant for a theory of language. A deterministic map between discrete linguistic representations like phonemes and continuous observables like articulatory movements would suggest that discrete representations can provide a more-or-less complete theory of human language, since continuous observables would be determined by the discrete representations. However, careful measurement of the continuous variables relevant for language complexifies this view. A variety of factors can gradiently but systematically modulate the relationship between discrete linguistic categories and their physical instantiation in articulation. For example, speech errors, while often subjectively perceived as categorical (e.g., Fromkin 1971), actually exhibit a gradient "trace" of the intended utterance in articulation (Mowrey & MacKay 1990; Pouplier & Goldstein 2010) and the resulting acoustics (Goldrick & Blumstein 2006; Alderete et al. 2021). If one accidentally says <cod> when they intended to say <god> (as in the example above), the errorful [k] will tend to have a voice onset time (VOT; primary phonetic dimension differentiating voiced and voiceless stops) which is slightly *shorter* (more similar to the voiced stop [g]) relative to an intended production of <cod>. In this case, the competitor <god> exerts a small, measurable influence on the articulation of <cod>. Lexical competitors also exert gradient effects on articulation in non-errorful speech. For instance, the VOT in <cod> is, on average, slightly longer than the VOT in <kid>, which has no minimal pair competitor *<gid> (e.g., Baese-Berk & Goldrick 2009; Wedel, Nelson & Sharp 2018). This effect, termed "contrastive hyperarticulation" (Wedel, Nelson & Sharp 2018), suggests a *dissimilatory* influence of lexical competitors on articulation, in the opposite direction as the *assimilatory* effect observed in speech errors. Social factors can also gradiently modulate speech articulation. For instance, speakers tend to gradiently shift their pronunciations towards those of their interlocutor—"converging" towards the interlocutor (e.g., Pardo, 2006)—an effect that is modulated by social characteristics of both the speaker and the listener like gender, race, and perceived attractiveness (e.g., Babel 2012; Wade 2022). During speech production, each of these (and other) influences interact in potentially complex ways. For instance, lexical competition in non-errorful speech interacts with the speaker's dialect (e.g., Clopper & Tamati 2014) and the perceived social identity of the interlocutor (Lee-Kim & Chou 2022).

The patterns described above demonstrate that the mapping from discrete linguistic representations to continuous articulatory movements is variable rather than deterministic. While some of this variability appears random (e.g., Whalen et al. 2018; Whalen & Chen 2019), other aspects of this variability are systematic, i.e., predicted by linguistic and social factors like those described above. Systematic variability in the relationship between discrete linguistic representations and continuous articulatory movements points to the need for a theory of this relationship. In other words, discrete representations, in combination with a deterministic map from discrete representations to continuous articulatory movements, *cannot* provide a complete theory of human language.

1.2 CONTINUOUS REPRESENTATIONS IN LINGUISTICS. A number of theories in linguistics formalize aspects of the relationship between discrete and continuous representations. For instance, in Articulatory Phonology (Browman & Goldstein 1986; Browman & Goldstein 1989; Browman &

Goldstein 1992), phonological representations consist of sets of constriction goals in the vocal tract, or *articulatory gestures*. Each gesture is constituted by a set of continuous variables like constriction location, constriction degree, and timing of gestural activation. Gestures drive physical movement of articulators through the Task Dynamic model (Saltzman & Munhall 1989). Even though gesture parameters are continuous representations which generate movement through continuous space, they also represent phonological contrast, since distinct parameters can represent distinct phonemes. For this reason, Articulatory Phonology has been argued to bridge the divide between discrete (phonological) and continuous (phonetic) representation (Iskarous 2017; Iskarous & Pouplier 2022). Articulatory Phonology is able to explain much of the systematic variability in the relationship between phonological representation and articulatory movement through temporal overlap of gestures and blending of gestural parameters.

In exemplar theory (Pierrehumbert 2001; Pierrehumbert 2002), discrete linguistic categories are constituted by sets of episodic memories or exemplars associated with that category. Each exemplar encodes detailed information on continuous phonetic dimensions. Each exemplar is also labeled as a member of one or more discrete categories, both linguistic and social. For instance, a single exemplar—containing detailed phonetic information—may simultaneously be associated with a particular phonological category, a particular lexical item, a particular person, and so on. In this way, linguistic representations contain both discrete and continuous information. While categories are inherently discrete, they exist in a shared continuous space, since they are constituted by (potentially overlapping) clouds of individual exemplars which are defined in this space. Exemplar theory has offered explanations for the relationship between word frequency and long-term phonetic change, as well as word-specific phonetic variation, for example (e.g., Wedel 2012; Hay et al. 2015).

Finally, connectionist or parallel distributed processing models (PDP: Rumelhart, McClelland, & PDP Research Group 1986; McClelland & Rogers 2003) represent categories as patterns of activation distributed over many interconnected processing units. The activation of each unit, and the strength of its connections to other units, are continuous variables which change during training with examples. While some individual units can represent categories, other units are not interpretable on their own. Rather, these units contribute to the representation of multiple categories through their activation and connections to other units. Dimensionality reduction methods like principal component analysis can reveal a continuous similarity structure which relates categories. PDP models offer explanations for phenomena related to category learning (e.g., Mareschal, French & Quinn 2000) and decay in aging or dementia (e.g., Murre, Graham & Hodges 2001), as well as phenomena related to real-time word recognition (e.g., Kawamoto, 1993). A particular framework in the PDP tradition, gradient symbol processing (GSP: Smolensky, Goldrick & Mathis 2014), explicitly relates distributed activation patterns to discrete representations from linguistic theory. By further linking these activation patterns to gestural timing in Task Dynamics, a GSP model of phonological planning derives the trace effect in speech errors as resulting from activation patterns that are intermediate in continuous similarity space between the intended and errorful category (Goldrick & Chu 2014).

The remainder of this paper is structured as follows. In Section 2, I provide a general introduction to Dynamic Field Theory (DFT: Schöner, Spencer & DFT Research Group 2016). DFT, when applied to human language, can be seen as part of the tradition of utilizing continuous representations in linguistic theory. In particular, Articulatory Phonology is expressed in the framework of dynamical systems (e.g., Port & Van Gelder, 1998), which is the same framework used by DFT, allowing a natural link between the theories. A relatively unique benefit of DFT is that it grounds representations in the dynamics of neural activity. In doing so, it provides tight constraints on possible representations and processes, and increases empirical falsifiability. In Section 3, I review research that has utilized DFT to bridge the gap between discrete and continuous aspects of linguistic representations. Section 4 concludes by outlining the general theory of language suggested by the research reviewed in Section 3.

2. Dynamic Field Theory. Dynamic Field Theory (DFT: Schöner, Spencer & DFT Research Group 2016) is a theoretical and modeling framework for understanding and simulating perception, action, and cognition as resulting from the coordinated activity of interconnected populations of neurons. In this section, I give a brief introduction to the framework. Readers who want to learn more are referred to the introductory textbook, Schöner et al. (2016). The textbook includes hands-on simulation exercises using the publicly available COSIVINA software package (Schneegans 2021). Other software packages for building and simulating DFT models, along with tutorials and other information, e.g. about an annual summer school, are available on the DFT website (dynamicfieldtheory.org).

DFT is motivated by the observation that cognition depends on an interplay between *stability* and *instability*, two technical concepts from dynamical systems theory (e.g., Kelso, 1995, 2012). Stability allows robustness of cognitive processes in the face of ubiquitous noise (e.g., neural and environmental), while instabilities allow qualitative transitions between transiently stable states, e.g., between a resting state and an active state, between one movement target and another, or between one percept and another. Another core motivation for DFT is the observation of *neural population representation*, i.e., that aggregated activity across a small set of neurons is a better predictor of behavior than the activity of any single cell (e.g., Georgopoulos, Schwartz & Kettner 1986; Jancke et al. 1999). Thus, a *mesoscopic* level of description (neural populations) is privileged relative to a microscopic (single cell) or macroscopic (whole brain region) level (Cohen & Newsome 2009; Schöner 2020). DFT formalizes neural population activity as a dynamical system which exhibits an interplay between stability and instability. In particular, excitatory connections between similarly tuned neurons and inhibitory connections between differently tuned neurons (Erlhagen et al. 1999; Jancke et al. 1999) allow transiently stable "peaks" of activation to form and dissipate as instabilities (Amari 1977).

2.1 NEURAL NODES. In DFT, both time and neural activation are represented as continuous variables. However, in some cases, a neural population represents a discrete category. For instance, some cells in medial temporal lobe respond selectively to stimuli related to particular people or objects like Jennifer Aniston or the Tower of Pisa (Quiroga et al. 2005; Quiroga 2012), and groups of cells in superior temporal gyrus respond selectively to sounds associated with particular phonetic categories (Mesgarani et al. 2014). In DFT, neural populations representing discrete categories are modeled as *neural nodes*. The activation dynamics of a neural node are given in Eq. 1:

$$\dot{u} = -u + h + s(t) + c \cdot g(u) + q \cdot \xi(t) \qquad (1)$$

Eq. 1 describes the rate of change of neural activation $\dot{u}$ as negatively related to current activation u. This relationship defines stable *point attractor* dynamics, in which activation is always attracted to a particular state in activation space. The position of that state is given by the sum of the other terms in the equation, which can change from moment to moment. h is the resting state; $s(t)$ is external input to the system, e.g., from another neural population, or from sensation; c is the strength of lateral connections within the neural population, weighted by a sigmoidal function of activation $g(u)$; and $\xi(t)$ is randomly distributed noise, weighted by q. While the $\dot{u} = -u$ component ensures point attractor dynamics, variation in the other terms over time changes the position of the point attractor, driving qualitative changes in behavior over time. Crucially, $c \cdot g(u)$ introduces *nonlinearity* into the system, such that when activation reaches a certain state, it can spontaneously jump to an even higher state due to lateral interaction. This spontaneous jump is an instability: in this case, a transition between a resting state and an active state. Instabilities allow discrete-seeming behavior, e.g., detection of a category, to arise in continuous time and continuous activation space.

2.2 NEURAL FIELDS. Some neurons, rather than responding to a discrete category, respond to a specific range of metric values on a continuous dimension, e.g., the direction of a manual reaching movement (Georgopoulos, Schwartz & Kettner 1986) or the position of a visual stimulus (Jancke et al. 1999). In the realm of speech, continuous acoustic properties of speech are encoded in superior temporal gyrus (Akbari et al. 2019), and continuous kinematic properties of articulatory movements are encoded in sensorimotor cortex (Chartier et al. 2018). In DFT, populations of neurons that are sensitive to the same continuous dimension are modeled as *neural fields*. The activation dynamics of a neural field are given in Eq. 2:

$$\dot{u}(x) = -u(x) + h + s(x,t) + \int k(x - x')g\big(u(x',t)\big)dx' + q \cdot \xi(x,t) \qquad (2)$$

Like neural nodes, neural fields are defined by point attractor dynamics in activation space, such that activation is always attracted to a particular state. Unlike neural nodes, activation in a neural field is defined over a continuous dimension x. Input s is also defined over x, such that input is localized to certain regions of the continuous dimension. Lateral interaction k, while still weighted by the sigmoidal function g, is defined over the metric distance between the neurons contributing activation x' and the neurons receiving activation x. For neurons that are sensitive to nearby regions of the continuous dimension (i.e., $x - x'$ is small), interaction k is excitatory. For neurons that are sensitive to more distant regions of the continuous dimension (i.e., $x - x'$ is large), interaction k is inhibitory. Like with neural nodes, lateral interaction in neural fields introduces nonlinearity which causes instabilities, i.e., spontaneous transitions between resting and active states. In the case of fields, active states are "peaks" of activation, localized to a specific region of the field. Local excitation stabilizes peaks of activation against noise and decay, and distal inhibition narrows the peak and prevents runaway activation. The location of an activation peak in the field contains information about the environment (in the case of a perception field) or about a movement goal (in the case of a motor field). For instance, a neural field representing the positions of objects in space may receive input $s(x,t)$ localized to a rightward region of the dimension x from a visual stimulus located to the right of the perceiver. This input may cause an activation peak on the rightward side of the field, representing detection of the object to the right of the receiver. Through synaptic projection to other neural populations, this activation peak may trigger peaks in other fields, or activate nodes representing categories, which may trigger further downstream processing through further synaptic connections. Simulated architectures containing interconnected nodes and fields can autonomously generate behavior, e.g., visually searching a scene (Grieben & Schöner 2022) or moving through a physical space (Bicho, Mallet & Schöner 2000).

2.3 RELATIONSHIP BETWEEN DFT AND NEURAL MEASUREMENTS. The basic neural assumptions underlying DFT are primarily motivated by single-cell recordings in animals (e.g., Georgopoulos, Schwartz & Kettner 1986; Jancke et al. 1999). For instance, neural fields have been constructed from the measured tuning curves of individual neurons in cat visual cortex (Erlhagen et al. 1999). Although it is possible to relate neural activation from DFT simulations to electrophysiological and hemodynamic recordings in humans (Wijeakumar et al. 2017), applications of DFT to humans have primarily been motivated and constrained by behavioral data (e.g., Erlhagen & Schöner 2002), under the assumption that the basic dynamics of neural population activity are similar across species. This is because single-neuron recording in humans is highly invasive, and simultaneous recording from many individual neurons in humans has so far been very difficult, although methods are improving, e.g., with the Neuropixels probe (Jun et al. 2017).

Although the aggregate activity in a neural population derives from the activity of the component neurons, the DFT formalisms of neural nodes and neural fields abstract away from the activity of each individual neuron. This allows for a low-dimensional description of neural activity which is directly relatable to cognitive, perceptual, and behavioral variables. However, the cost of such a description is that some details are lost. For instance, individual neural spikes are not modeled. Rather, aggregated spiking activity across a neural population is estimated via integration over short timescales. This estimation or "mean-field approximation" relies on particular assumptions, e.g., that spiking activity across a population is sufficiently frequent and asynchronous, which does not hold in all cases (e.g., Faugeras, Touboul & Cessac 2009). For instance, synchronized oscillatory activity in neural populations has been argued to be important for some cognitive processes (e.g., Churchland et al. 2012), including some related to speech (e.g., Poeppel & Assaneo 2020). Moreover, some processes related to long-term change in synaptic strength depend on the temporal phasing between pre- and post-synaptic spikes (e.g., Markram et al. 1997). These kinds of processes represent instances where the assumptions of DFT break down. Luckily, DFT uses a very general formalism for describing neural activity: differential equations. Thus, there is no obvious reason why the low-dimensional neural dynamics of DFT could not, in principle, be explicitly coupled to higher-dimensional neural processes involving individual spikes—also expressible in differential equations (e.g., Eliasmith 2013)—in cases where this greater level of detail is found to be necessary for understanding the relevant cognitive process.

Finally, it is important to note that the researcher using DFT must hypothesize which categories are represented by nodes and which continuous dimensions are represented by fields. With respect to DFT models of language, this leaves a crucial role for theoretical linguistics, which has unique insight into the categories and dimensions relevant for language. This approach contrasts with a "neural manifold" approach in which dimensions are derived directly from neural activation data (Langdon, Genkin & Engel 2023). In some cases, dimensions motivated by theory have been found to coincide with dimensions of representation derived from measured neural activity, e.g., movement direction in fruit flies (Seelig & Jayaraman 2015) and mice (Chaudhuri et al. 2019). As methods improve for simultaneously recording from large numbers of individual cells in humans, researchers from linguistics and neuroscience may be able to use both perspectives to triangulate the dimensions relevant for the neural dynamics of language.

3. Dynamic Field Theory in linguistics. Since the inception of DFT in the early 1990s, most work in the framework has focused on aspects of action, perception, and cognition outside of language per se. However, in the last five years or so, a growing body of work has applied DFT to understand a variety of phenomena related to speech and language.[1] The continuous dimensions of neural fields, as well as the nonlinear activation dynamics giving rise to instabilities, have proven particularly useful in unifying discrete and continuous aspects of linguistic representations. This section is divided into subsections based on the primary subfield relevant to each model; however, due to the inherent potential for unification across DFT models (discussed further in Section 4), each model also has implications outside of its primary subfield. My aim in this section is to give a brief description of each model so that interested readers can refer to the original papers to learn more.

3.1 SPEECH PRODUCTION. One of the first models of speech production in the framework of DFT was proposed by Roon & Gafos (2016). The model simulates a response-distractor task in which participants are visually cued to produce a particular syllable in the presence of an auditory distractor. Key findings from this paradigm are that participants are slower to start speaking when the distractor mismatches the target in a phonetic feature (Galantucci, Fowler & Goldstein 2009; Roon & Gafos 2015), and that articulation is gradiently pulled in movement directions associated with the distractor (Yuen et al. 2010). Roon & Gafos (2016) propose an architecture of coupled neural fields governing the constriction location and constriction degree targets for various articulators (lower lip, tongue tip, tongue back) as well as VOT. The intention to produce a particular syllable is modeled as a set of field inputs corresponding to the target values necessary to produce that syllable. For instance, when the simulated participant is cued to produce /t/, inputs influence the tongue tip fields towards a small constriction degree (full closure) located at the alveolar ridge, and an input influences the VOT field towards a high VOT (voiceless). When activation peaks stabilize in all of the fields, the locations of the peaks are assumed to determine the targets of articulatory movement. Moreover, the time it takes for activation peaks to form is taken as a proxy for response time in the task. Crucially, input from the perceived distractor also influences the same neural fields (for a review of evidence for interaction between perception and production, see Fowler 2016). This derives the two key results from the experimental task in the following way. When an input from the perceptual distractor mismatches an input for the target, lateral inhibition between the two activated field regions slows down the formation of an activation peak, leading to a slowed response time. Moreover, the distractor input can exert a gradient attractive influence on the location of the activation peak, exerting a gradient influence on the target of articulatory movement. Note that, in the model (and in the models described below), differences between discrete categories (e.g., /t/ vs. /d/) are modeled as continuous differences between field input distributions (e.g., centered on a high vs. a low VOT value). The nonlinear dynamics of lateral neural interaction allows the discrete selection of one category over another on each trial. At the same time, temporal continuity allows the derivation of gradient response time predictions, and feature continuity allows the derivation of gradient articulatory predictions. In this way, discreteness and continuity coexist.

---

[1] Some models in linguistics have utilized continuous fields of possible gestural target values, but are otherwise outside of the neural framework of DFT. For instance, Tilsen (2019) and Kim & Tilsen (2025) blend gestural intentions—modeled as distributions—through summation, in order to model effects related to simultaneously active supralaryngeal (Tilsen 2019) or laryngeal (Kim & Tilsen 2025) gestures; Gafos & Kirov (2009) model long-term change in phonetic representations via updating of distributions representing long-term memories of phonetic targets. Since these models do not incorporate the nonlinear neural dynamics of DFT that give rise to instabilities through lateral interaction, they are outside the scope of this review.

A number of subsequent studies have explored the potential for input interaction in speech target fields to explain a variety of phenomena. For instance, Stern et al. (2022) model the trace effect in voicing errors (see Section 1.1) as resulting from the interaction of inputs to a VOT planning field. In this model, two discrete inputs corresponding to the intended and errorful intentions can lead to a single activation peak that is intermediate between the two intentions. Stern & Shaw (2023a) extend this model to capture the contrastive hyperarticulation effect in non-errors by introducing an *inhibitory* input (cf. Tilsen 2019) corresponding to the minimal pair competitor which repels the activation peak away from the competitor. In this way, both trace effects in errors and contrastive hyperarticulation in non-errors can be understood as arising from variation in a single parameter: the polarity of input (excitatory or inhibitory) from a competitor category. Stern & Shaw (2023b) show how input polarity can be derived (rather than stipulated) from the coupling dynamics between neural fields representing articulatory targets (voicing, constriction location, constriction degree) and neural nodes representing lexical items. This model demonstrates that response times and articulatory targets are affected not only by the number of phonological neighbors in the lexicon (neighborhood density), but also by the particular way that those neighbors are coupled to overlapping phonetic dimensions, which may underlie the apparently conflicting set of results bearing on phonological neighborhood density effects. Finally, Pintado-Urbanc (2025, this issue) models gradient effects of bilingual code-switching on VOT via the interaction of inputs from the two languages of a bilingual speaker. In this model, the distinct languages of a bilingual speaker interact in a shared neural field representing VOT targets. Moreover, dominance-based asymmetries in the phonetic effects of code-switching (Olson 2013) arise from differences in the strength of input from each language based on relative frequency of use, by analogy with the inverse frequency effect in syntactic priming (e.g., Ferreira 2003).

3.2 PHONOLOGY. Other models have linked the real-time neural dynamics of speech production to qualitative patterns typically addressed in the subfield of phonology. Chaturvedi & Shaw (2025) model tonal downstep in Bimoba, a Gur language spoken in Ghana (Snider 1998), through the interaction of inputs to a neural field governing pitch targets. In particular, a combination of inputs corresponding to high tone, low tone, and low register, in a particular temporal sequence, derive the phonological downstep pattern. In the model, the temporal sequencing of tones is implemented via a *condition of satisfaction* mechanism (Sandamirskaya & Schöner 2008), in which perception of an achieved target triggers inhibition of the field location associated with that target, and excitation of the field location associated with the next target, facilitating rapid transitions between movement targets. Shaw (2025) presents a neural field model governing glottal width targets which derives a phonological typology of voicing in /CVCV/ sequences, where both consonants are underlyingly voiceless. Given the same set of inputs to the neural field, three distinct output patterns are possible—faithful voicing contour, vowel devoicing, or intervocalic voicing—depending on the timing of inputs to the field. In this model, input timing is derived from an oscillator-based clock representing moras. Both Chaturvedi & Shaw (2025) and Shaw (2025) represent two kinds of advancements in DFT modeling of speech and language. First, neural field activation dynamics are explicitly coupled to articulatory movement dynamics (cf. Kirkham & Strycharczuk 2024; Kim & Tilsen 2025), so that neural fields autonomously drive movement of articulators. Second, the timing of field inputs is derived, rather than stipulated: in Chaturvedi & Shaw (2025), via the condition of satisfaction; in Shaw (2025), via an oscillator-based timing mechanism. Both models demonstrate how the gradient temporal details of activation peak stabilization and dissipation can lead to categorically distinct patterns in articulation, typically described as phonological, despite the underlyingly continuous state space.

Other DFT models have addressed change in phonological representations over a timescale longer than individual utterances. Shaw & Tang (2023) model the phenomenon of “leaky prosody” (Tang & Shaw 2021), whereby words exhibit phonetic characteristics of the prosodic contexts in which they are typically produced, even when produced in other contexts. In the model, a neural field governing pitch targets is simultaneously influenced by inputs from lexical, tonal, and prosodic intentions, which jointly determine the location of the activation peak (pitch target) during a given utterance. Each time an activation peak forms, it contributes a small update to the field location of the lexical input (cf. Gafos & Kirov 2009), i.e., the long-term phonetic memory of the word (Pierrehumbert 2002). Since each activation peak bears some influence from the prosodic input, each lexical input takes on characteristics of the prosodic inputs that it tends to co-occur with. In this way, the content of categorical knowledge (inputs corresponding to lexical items) changes gradiently over a slow timescale as the result of many individual utterance events. Kirkham & Strycharczuk (2024) utilize a similar input-updating mechanism to model a long-term phonological change from a vowel monophthong to a diphthong. In their model, two inputs influence a neural field governing tongue body constriction location. The timing of the inputs is controlled using an oscillator, and activation peaks drive articulatory movement via coupling between neural and articulatory dynamics (cf. Chaturvedi & Shaw 2025; Shaw 2025). To produce a monophthong, both inputs have similar locations in the field, so that there is no transition between vowel targets. To produce a diphthong, each input has a different location in the field so that the tongue body moves to a different location when the second input creates an activation peak. To model long-term change from a monophthong to a diphthong, they simulate two speakers interacting with each other, one with a set of inputs corresponding to a monophthong, and one with a set of inputs corresponding to a diphthong. Each time that the monophthongal speaker perceives a diphthongal target, their input distributions update to be slightly more diphthongal, driving long-term convergence towards the diphthongal speaker.

3.3 LEXICAL SEMANTICS. Some work in the DFT framework has addressed the lexical semantics of spatial relation terms like *above* or *to the left of*, as well as perceptual descriptors like *red* or *big*. Richter et al. (2021) model the meaning of each term as a coupling relation between a neural node representing the lexical item, and neural fields representing relevant continuous dimensions. For instance, the node corresponding to *above* is coupled to the upper region of a neural field representing the spatial relation between objects, and the node corresponding to *red* is coupled to the red region of a neural field representing the spectrum of visible color. The model, consisting of a network of coupled nodes and fields, is able to autonomously search a visual scene for a target object. Stern & Piñango (2024) extend the notion of lexical meaning as node-field coupling to semantic dimensions underlying lexical polysemy. In particular, polysemy is understood as arising from a wide distribution in the coupling pattern, such that the same lexical item can elicit a wide range of possible interpretations along the relevant dimensions. Using the English lexical item *have* as a test case, this model captures the influence of preceding context on lexical interpretation via the persistence through time of neural field activation states. Bhat, Spencer & Samuelson (2022) combine the real-time dynamics of word interpretation with the long-term dynamics of word learning in the Word-Object Learning via Visual Exploration in Space (WOLVES) model. As the model autonomously searches a visual scene, it also learns to associate attended objects with perceived words via the strengthening of synaptic connections that couple these representations to each other.

3.4 SYNTAX. Sabinasz, Richter & Schöner (2023) extend the model of Richter et al. (2021) to nested noun phrases which combine multiple lexical items in a syntactic structure like *the small tree to the left of the big tree* or *the big tree to the left of the lake and to the right of the house*. This model offers a solution to the "problem of 2"—in which a single concept has multiple simultaneous but distinct instantiations—and the "massiveness of the binding problem"—in which a single concept can participate in multiple relations (Jackendoff 2002)—through "index" fields which dynamically bind lexical items to each other. Kati et al. (2024) introduce another set of extensions to this model in order to address negation and truth value judgment. By adding a network of six coupled neural nodes, this model is able to evaluate the truth of true affirmative phrases, false affirmative phrases, true negated phrases, and false negated phrases. The model also mirrors empirical patterns in human response times across these conditions.

4. Towards a neural dynamic theory of language. The models reviewed in Section 3, taken together, represent a unified perspective on human language. Language behavior is constituted by activation patterns across neural populations which drive production and comprehension; the dynamics of these activation patterns are formalized in differential equations. Linguistic knowledge is constituted by patterns of synaptic connectivity within and between neural populations; synaptic connections are formalized as coupling between the differential equations describing neural activation. Since DFT uses only one language of description—differential equations—there is no need for "translation" across "interfaces". Rather, language is a single dynamical system formalizable as a set of coupled differential equations. The system is defined in a continuous state space, but it exhibits discrete, categorical behavior because of nonlinearity in the dynamics. The models reviewed above each contain different components because they were built to address a particular set of phenomena in language. However, each of these models should be viewed as a component of a single architecture of language. This architecture can really be built, piece by piece, by slotting models together through coupling of the relevant differential equations. For instance, the WOLVES model (Bhat, Spencer & Samuelson 2022) is the result of slotting together the word-object learning model (WOL: Samuelson et al. 2011) and the visual exploration in space model (VES: Schneegans, Spencer & Schöner 2016) in this way. Ultimately, a DFT architecture for language would be simultaneously a model of processing, knowledge, and learning. Each of these empirical domains can be studied as different timescales of evolution in the same model, which is a model of a single brain. As a final point, while I have emphasized here the continuous representations of neural fields, the general success of research in linguistics to date suggests that discrete representations are a crucial component of linguistic knowledge and processing. As we continue to build and refine a neural dynamic architecture of language, we are likely to find that neural nodes, coupled to each other through synaptic projection, play a fundamental role. However, we are also likely to find that continuous representations play important roles in domains previously conceptualized as categorical, like lexical semantics (e.g., Stern & Piñango 2024).

References

Akbari, Hassan, Bahar Khalighinejad, Jose L. Herrero, Ashesh D. Mehta & Nima Mesgarani. 2019. Towards reconstructing intelligible speech from the human auditory cortex. *Scientific Reports* 9(1). 874. http://doi.org/10.1038/s41598-018-37359-z.

Alderete, John, Melissa Baese-Berk, Keith Leung & Matthew Goldrick. 2021. Cascading activation in phonological planning and articulation: Evidence from spontaneous speech errors. *Cognition* 210. http://doi.org/10.1016/j.cognition.2020.104577.

Amari, Shun-ichi. 1977. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics* 27(2). 77–87. doi.org/10.1007/BF00337259.

Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40(1). 177–189. doi.org/10.1016/j.wocn.2011.09.001.

Baese-Berk, Melissa & Matthew Goldrick. 2009. Mechanisms of interaction in speech production. *Language and Cognitive Processes* 24(4). 527–554.

Bhat, Ajaz A., John P. Spencer & Larissa K. Samuelson. 2022. Word-Object Learning via Visual Exploration in Space (WOLVES): A neural process model of cross-situational word learning. *Psychological Review* 129(4). 640–695. doi.org/10.1037/rev0000313.

Bicho, Estela, Pierre Mallet & Gregor Schöner. 2000. Target representation on an autonomous vehicle with low-level sensors. *The International Journal of Robotics Research* 19(5). 424–447. doi.org/10.1177/02783640022066950.

Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6(2). 201–251. doi.org/10.1017/S0952675700001019.

Browman, Catherine P. & Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49(3–4). 155–180. doi.org/10.1159/000261913.

Browman, Catherine P. & Louis M. Goldstein. 1986. Towards an articulatory phonology. *Phonology* 3. 219–252. doi.org/10.1017/S0952675700000658.

Chartier, Josh, Gopala K. Anumanchipalli, Keith Johnson & Edward F. Chang. 2018. Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex. *Neuron* 98(5). 1042-1054.e4. doi.org/10.1016/j.neuron.2018.04.031.

Chaturvedi, Manasvi & Jason A. Shaw. 2025. A dynamic neural model of tonal downstep. *Proceedings of the Annual Meetings on Phonology* 1(1). doi:10.7275/amphonology.3021.

Chaudhuri, Rishidev, Berk Gerçek, Biraj Pandey, Adrien Peyrache & Ila Fiete. 2019. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature Neuroscience* 22(9). 1512–1520.

Churchland, MM, JP Cunningham, MT Kaufman, JD Foster, P Nuyujukian, SI Ryu & KV Shenoy. 2012. Neural population dynamics during reaching. *Nature* 487(7405). 51–56.

Clopper, Cynthia G. & Terrin N. Tamati. 2014. Effects of local lexical competition and regional dialect on vowel production. *The Journal of the Acoustical Society of America* 136. 1–4.

Cohen, Marlene R. & William T. Newsome. 2009. Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *Journal of Neuroscience* 29(20). 6635–6648. doi.org/10.1523/JNEUROSCI.5179-08.2009.

Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93(3). 283–321. doi.org/10.1037/0033-295X.93.3.283.

Eliasmith, Chris. 2013. *How to build a brain: A neural architecture for biological cognition*. Oxford: Oxford University Press.

Erlhagen, Wolfram, Annette Bastian, Dirk Jancke, Alexa Riehle & Gregor Schöner. 1999. The distribution of neuronal population activation (DPA) as a tool to study interaction and integration in cortical representations. *Journal of Neuroscience Methods* 94(1). 53–66.

Erlhagen, Wolfram & Gregor Schöner. 2002. Dynamic field theory of movement preparation. *Psychological Review* 109(3). 545–572. doi.org/10.1037/0033-295X.109.3.545.

Faugeras, Olivier D., Jonathan D. Touboul & Bruno Cessac. 2009. A constructive mean-field analysis of multi population neural networks with random synaptic weights and stochastic inputs. *Frontiers in Computational Neuroscience* 3.

Ferreira, Victor S. 2003. The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language* 48(2). 379–398.

Fowler, Carol A. 2016. Speech perception as a perceptuo-motor skill. In Gregory Hickok & Steven L. Small (eds.), *Neurobiology of language*. 175–184. San Diego: Academic Press.

Fromkin, Victoria A. 1971. The non-anomalous nature of anomalous utterances. *Language* 47(1). 27–52.

Gafos, Adamantios I. & Christo Kirov. 2009. A dynamical model of change in phonological representations: The case of lenition. *Approaches to Phonological Complexity*, 219–240.

Gahl, Susanne, Yao Yao & Keith Johnson. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language* 66(4). 789–806. doi.org/10.1016/j.jml.2011.11.006.

Galantucci, B., C.A. Fowler & L. Goldstein. 2009. Perceptuomotor compatibility effects in speech. *Attention, Perception, and Psychophysics* 71(5). 1138–1149.

Georgopoulos, A. P., A. B. Schwartz & R. E. Kettner. 1986. Neuronal population coding of movement direction. *Science* 233. 1416–1419.

Goldrick, Matthew & Sheila Blumstein. 2006. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes* 21(6). 649–683. doi.org/10.1080/01690960500181332.

Goldrick, Matthew & Karen Chu. 2014. Gradient co-activation and speech error articulation: Comment on Pouplier and Goldstein (2010). *Language, Cognition and Neuroscience* 29(4). 452–458. doi.org/10.1080/01690965.2013.807347.

Grieben, Raul & Gregor Schöner. 2022. Bridging DFT and DNNs: A neural dynamic process model of scene representation, guided visual search and scene grammar in natural scenes. *Proceedings of the Annual Meeting of the Cognitive Science Society* 44.

Hay, Jennifer B., Janet B. Pierrehumbert, Abby J. Walker & Patrick LaShell. 2015. Tracking word frequency effects through 130 years of sound change. *Cognition* 139. 83–91.

Iskarous, Khalil. 2017. The relation between the continuous and the discrete: A note on the first principles of speech dynamics. *Journal of Phonetics* 64. 8–20.

Iskarous, Khalil & Marianne Pouplier. 2022. Advancements of phonetics in the 21st century: A critical appraisal of time and space in Articulatory Phonology. *Journal of Phonetics* 95. 101195. doi.org/10.1016/j.wocn.2022.101195.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Jancke, Dirk, Wolfram Erlhagen, Hubert R. Dinse, Amir C. Akhavan, Martin Giese, Axel Steinhage & Gregor Schöner. 1999. Parametric population representation of retinal location: Neuronal interaction dynamics in cat primary visual cortex. *Journal of Neuroscience* 19(20). 9016–9028. doi.org/10.1523/jneurosci.19-20-09016.1999.

Jun, James J., Nicholas A. Steinmetz, Joshua H. Siegle, Daniel J. Denman, Marius Bauza, Brian Barbarits, Albert K. Lee, et al. 2017. Fully integrated silicon probes for high-density recording of neural activity. *Nature* 551(7679). 232–236. doi.org/10.1038/nature24636.

Kati, Lea, Daniel Sabinasz, Gregor Schöner & Barbara Kaup. 2024. Interaction of polarity and truth value - A neural dynamic architecture of negation processing. *Proceedings of the Annual Meeting of the Cognitive Science Society* 46.

Kawamoto, Alan H. 1993. Nonlinear dynamics in the resolution of lexical ambiguity: A parallel distributed processing account. *Journal of Memory and Language* 32(4). 474–516.

Kelso, J. A. Scott. 1995. *Dynamic patterns: The self-organization of brain and behavior*. Cambridge: MIT Press.

Kelso, J. A. Scott. 2012. Multistability and metastability: Understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367(1591). 906–918. doi.org/10.1098/rstb.2011.0351.

Kim, Seung-Eun & Sam Tilsen. 2025. The Gesture-Field-Register (GFR) framework for modeling F0 control. *Journal of Phonetics* 110. 101410. doi.org/10.1016/j.wocn.2025.101410.

Kirkham, Sam & Patrycja Strycharczuk. 2024. A dynamic neural field model of vowel diphthongisation. *Proceedings of the 13th International Seminar on Speech Production*.

Langdon, Christopher, Mikhail Genkin & Tatiana A. Engel. 2023. A unifying perspective on neural manifolds and circuits for cognition. *Nature Reviews Neuroscience* 24(6). 363–377. doi.org/10.1038/s41583-023-00693-x.

Lee-Kim, Sang-Im & Yun-Chieh Iris Chou. 2022. Unmerging the sibilant merger among speakers of Taiwan Mandarin. *Laboratory Phonology* 13(1). doi.org/10.16995/labphon.6446.

Levelt, Willem J.M., Ardi Roelofs & Antje S. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22(1). 1–75.

Mareschal, Denis, Robert M. French & Paul C. Quinn. 2000. A connectionist account of asymmetric category learning in early infancy. *Developmental Psychology* 36(5). 635–645.

Markram, Henry, Joachim Lübke, Michael Frotscher & Bert Sakmann. 1997. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* 275(5297). 213–215. doi.org/10.1126/science.275.5297.213.

McClelland, James L. & Timothy T. Rogers. 2003. The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience* 4(4). 310–322.

Mesgarani, Nima, Connie Cheung, Keith Johnson & Edward F. Chang. 2014. Phonetic feature encoding in human superior temporal gyrus. *Science* 343(6174). 1006–1010.

Mowrey, Richard A. & Ian R. A. MacKay. 1990. Phonological primitives: Electromyographic speech error evidence. *The Journal of the Acoustical Society of America* 88(3). 1299–1312. doi.org/10.1121/1.399706.

Murre, Jaap M. J., Kim S. Graham & John R. Hodges. 2001. Semantic dementia: Relevance to connectionist models of long-term memory. *Brain* 124(4). 647–675.

Olson, Daniel J. 2013. Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics* 41(6). 407–420.

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4). 2382–2393. doi.org/10.1121/1.2178720.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*. 137–158. Amsterdam: John Benjamins.

Pierrehumbert, Janet B. 2002. Word-specific phonetics. *Laboratory Phonology* 7. 101–140. Berlin: Mouton de Gruyter.

Pintado-Urbanc, Alessandra. 2025. A dynamic neural field model of asymmetric interference effects in code-switching. *Proceedings of the Linguistic Society of America*, vol. 10.

Poeppel, David & M. Florencia Assaneo. 2020. Speech rhythms and their neural foundations. *Nature Reviews. Neuroscience* 21(6). 322–334. doi.org/10.1038/s41583-020-0304-4.

Port, Robert F. & Timothy Van Gelder (eds.). 1998. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge: The MIT Press.

Pouplier, Marianne & Louis Goldstein. 2010. Intention in articulation: Articulatory timing in alternating consonant sequences and its implications for models of speech production. *Language and Cognitive Processes* 25(5). 616–649.

Quiroga, R. Quian, L. Reddy, G. Kreiman, C. Koch & I. Fried. 2005. Invariant visual representation by single neurons in the human brain. *Nature* 435(7045). 1102–1107.

Quiroga, Rodrigo Quian. 2012. Concept cells: the building blocks of declarative memory functions. *Nature Reviews Neuroscience* 13(8). 587–597. doi.org/10.1038/nrn3251.

Richter, Mathis, Jonas Lins & Gregor Schöner. 2021. A neural dynamic model of the perceptual grounding of spatial and movement relations. *Cognitive Science* 45(10). e13045.

Roon, Kevin D. & Adamantios I. Gafos. 2015. Perceptuo-motor effects of response-distractor compatibility in speech: Beyond phonemic identity. *Psychonomic Bulletin & Review* 22(1). 242–250. doi.org/10.3758/s13423-014-0666-6.

Roon, Kevin D. & Adamantios I. Gafos. 2016. Perceiving while producing: Modeling the dynamics of phonological planning. *Journal of Memory and Language* 89. 222–243.

Rumelhart, David E., James L. McClelland, & PDP Research Group. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*. Cambridge: The MIT Press.

Sabinasz, Daniel, Mathis Richter & Gregor Schöner. 2023. Neural dynamic foundations of a theory of higher cognition: The case of grounding nested phrases. *Cognitive Neurodynamics*. 1–23.

Saltzman, Elliot L. & Kevin G. Munhall. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1(4). 333–382.

Samuelson, Larissa K., Linda B. Smith, Lynn K. Perry & John P. Spencer. 2011. Grounding word learning in space. *PLOS ONE* 6(12). doi.org/10.1371/journal.pone.0028095.

Sandamirskaya, Yulia & Gregor Schöner. 2008. Dynamic field theory of sequential action: A model and its implementation on an embodied agent. *7th IEEE International Conference on Development and Learning*. Monterey, CA: IEEE. 133–138.

Schneegans, Sebastian. 2021. COSIVINA: A Matlab toolbox to compose, simulate, and visualize neurodynamic architectures. https://github.com/cosivina/cosivina.

Schneegans, Sebastian, John P. Spencer & Gregor Schöner. 2016. Integrating "what" and "where": Visual working memory for objects in a scene. *Dynamic thinking: A primer on dynamic field theory*, 197–226. Oxford: Oxford University Press.

Schöner, Gregor. 2020. The dynamics of neural populations capture the laws of the mind. *Topics in Cognitive Science* 12(4). 1257–1271. doi.org/10.1111/tops.12453.

Schöner, Gregor, John Spencer & DFT Research Group. 2016. *Dynamic Thinking: A Primer on Dynamic Field Theory*. Oxford: Oxford University Press.

Seelig, Johannes D. & Vivek Jayaraman. 2015. Neural dynamics for landmark orientation and angular path integration. *Nature* 521(7551). 186–191. doi.org/10.1038/nature14446.

Shaw, Jason A. 2025. Unifying phonological and phonetic aspects of speech in dynamic neural fields: The case of laryngeal patterns in Japanese. *Phonological Studies* 28. 1–12.

Shaw, Jason A. & Kevin Tang. 2023. A dynamic neural field model of leaky prosody: Proof of concept. *Proceedings of the 2022 Annual Meeting on Phonology (AMP)*.

Smolensky, Paul, Matthew Goldrick & Donald Mathis. 2014. Optimization and quantization in gradient symbol systems: A framework for integrating the continuous and the discrete in cognition. *Cognitive Science* 38(6). 1102–1138. doi.org/10.1111/cogs.12047.

Snider, Keith L. 1998. Phonetic realisation of downstep in Bimoba. *Phonology* 15(1). 77–101. doi.org/10.1017/S0952675798003534.

Stern, Michael C., Manasvi Chaturvedi & Jason A. Shaw. 2022. A dynamic neural field model of phonetic trace effects in speech errors. *Proceedings of the Annual Meeting of the Cognitive Science Society* 44.

Stern, Michael C. & Maria M. Piñango. 2024. Contextual modulation of language comprehension in a dynamic neural model of lexical meaning. Manuscript. doi.org/10.48550/arXiv.2407.14701.

Stern, Michael C. & Jason A. Shaw. 2023a. Neural inhibition during speech planning contributes to contrastive hyperarticulation. *Journal of Memory and Language* 132. 104443. doi.org/10.1016/j.jml.2023.104443.

Stern, Michael C. & Jason A. Shaw. 2023b. Not all phonological neighbors affect production equivalently: Predictions from a neural dynamic model. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*. 4002–4006.

Tang, Kevin & Jason A. Shaw. 2021. Prosody leaks into the memories of words. *Cognition* 210. doi.org/10.1016/j.cognition.2021.104601.

Tilsen, Sam. 2019. Motoric mechanisms for the emergence of non-local phonological patterns. *Frontiers in Psychology* 10. 1–25. doi.org/10.3389/fpsyg.2019.02143.

Wade, Lacey. 2022. Experimental evidence for expectation-driven linguistic convergence. *Language* 98(1). doi.org/10.1353/lan.0.0257.

Wedel, Andrew. 2012. Lexical contrast maintenance and the organization of sublexical contrast systems. *Language and Cognition* 4(4). 319–355. doi.org/10.1515/langcog-2012-0018.

Wedel, Andrew, Noah Nelson & Rebecca Sharp. 2018. The phonetic specificity of contrastive hyperarticulation in natural speech. *Journal of Memory and Language* 100. 61–88. doi.org/10.1016/j.jml.2018.01.001.

Whalen, D. H., Wei Rong Chen, Mark K. Tiede & Hosung Nam. 2018. Variability of articulator positions and formants across nine English vowels. *Journal of Phonetics* 68. 1–14. doi.org/10.1016/j.wocn.2018.01.003.

Whalen, D. H. & Wei-Rong Chen. 2019. Variability and central tendencies in speech production. *Frontiers in Communication* 4. doi.org/10.3389/fcomm.2019.00049.

Wijeakumar, Sobanawartiny, Joseph P. Ambrose, John P. Spencer & Rodica Curtu. 2017. Model-based functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach. *Journal of Mathematical Psychology* 76. 212–235.

Wright, Richard. 2004. Factors of lexical competition in vowel articulation. In J. Local, R. Ogden & R. Temple (eds.), *Papers in Laboratory Phonology VI*, 75–87. Cambridge: Cambridge University Press. doi.org/10.1017/cbo9780511486425.005.

Yuen, Ivan, Matthew H. Davis, Marc Brysbaert & Kathleen Rastle. 2010. Activation of articulatory information in speech perception. *Proceedings of the National Academy of Sciences of the United States of America* 107(2). 592–597. doi.org/10.1073/pnas.0904774107.