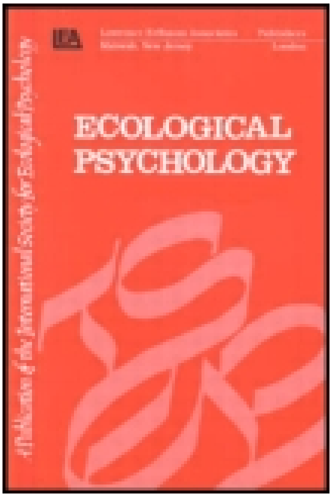

Ecological Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/heco20

An Integrative Research Strategy for Exploring Synergies in Natural Language Performance

Rick Dale[a]
[a] Cognitive and Information Sciences University of California, Merced
Published online: 14 Aug 2015.

Ecological Psychology, 27:190–201, 2015

ISSN: 1040-7413 print/1532-6969 online
DOI: 10.1080/10407413.2015.1068649

An Integrative Research Strategy for Exploring Synergies in Natural Language Performance

Rick Dale
Cognitive and Information Sciences
University of California, Merced

I argue that a synthetic explanatory strategy is an important step forward for developing a more comprehensive understanding of human language, in particular natural language performance. By considering some statistical definitions of the notion of synergy, I argue that we should expect very low-level units of measure, such as voice or body movement, to predict higher level units of measure, such as complex linguistic functions like discourse. It would be useful to tie the structured behavior at this low level to high-level constraints, such as discourse goals. I refer to this as an "integrative strategy" to getting the synthetic approach going, and I briefly sketch our preliminary methods in conducting this integration across different timescales.

In a thorough treatment of human sociality and interaction, Enfield (2013) described contributions to conversation as unified composites: "A typically multimodal, multidimensional utterance will consist of numerous signs in a unified composite, e.g., words and morphemes, some morphosyntactic arrangement of these, some configuration of the hand, some movement of the arm in a certain direction and at a certain speed, some deployment of the artifactual environment, and much more besides" (p. 65). This is the natural context of human interaction. It is the context in which we learn our first language. It is richly structured across many timescales, and these scales, somehow, coalesce gracefully when we interact.

Correspondence should be addressed to Rick Dale, Cognitive and Information Sciences, University of California, Merced, Merced, CA 95343. E-mail: rdale@merced.edu

How this coalescence happens is still relatively mysterious. This mystery is a result, in part, of our tendency to isolate behaviors of interest in our sciences of language. This is justifiable, and it's a productive research strategy, given the complexity we face. But it is important that natural linguistic interaction exploits many modalities and dimensions simultaneously, the way Enfield (2013) described. It is also theoretically nontrivial that these modalities—nested levels of structure from sounds to discourse goals—are coordinated so coherently within and between people. It calls for a synthetic research strategy in addition to the more common analytic one.

This article is a brief review of this problem of synthesis from the perspective of synergies. I argue that the concept of a synergy may be useful to pursue solutions to this problem, despite the fact that quantifying synergies in Enfield's (2013) sense requires multiple and quite different measurement scales. As I describe later, this is a strategy distinct from the one commonly conducted in which we identify the functional relationship, say, between vision and social understanding (D. C. Richardson, Dale, & Tomlinson, 2009), phonetic variation and discourse context (Lindblom, 1990), and so on. I argue that a system-level characterization of the integration of behaviors will be useful to taking the next step to trace the coalescence of all these measures. In the tradition of interaction-dominant studies of human cognition, the integrative approach I sketch would characterize the coordination dynamics of the one- or two-person system under study (e.g., Riley, Richardson, Shockley, & Ramenzoni, 2011; Van Orden, Holden, & Turvey, 2003).

In what follows, I first selectively recap the concept of a synergy for the purpose of this discussion. Following this, I describe an "integrative strategy," starting from these basic principles. An explicitly quantitative definition of synergy recommends a broad data-driven strategy that, we hope, can approach this problem using tools from both the computational and cognitive sciences.[1] To conclude, I argue that synergies may also provide bridge principles for connecting an ecological approach to language to other cognitive approaches and thus support a more theoretically synthetic approach to natural language use.

WHAT IS A SYNERGY?

The readers of this special issue are likely familiar with this concept. Its thrust is relatively easy to convey: "The term 'synergy' is sometimes used to refer to

[1] I deliberately jump between the singular and plural pronouns because the empirical work, which I cannot describe in detail here, is still in an ongoing phase with several excellent researchers, including Alexandra Paxton, Shreya Gupta, Pooja Patel, and Brittany Oakes. The use of "we" indicates the teamwork where appropriate.

softly assembled systems—a functional grouping of structural elements (molecules, genes, neurons, muscles, limbs, individuals, etc.) that are temporarily constrained to act as a single coherent unit" (M. J. Richardson, Dale, & Marsh, 2014, p. 254). This casual definition echoes the synthetic concern described earlier. The explanatory goal of those who employ this term, and related notions, is to understand how a system is integrated through several (or perhaps many) interdependent parts.[2]

As readers also no doubt know, the notion of synergy has been especially productive in the study of motor control (Bernstein, 1967; Kelso, 2009). The many muscles involved in an organized action have a large number of potential degrees of freedom; their effective degrees of freedom, however, are of much lower dimensionality. The effective degrees of freedom are constrained through the interdependence of the muscles and joints interacting during the performance. But how would we discern this quantitatively? The statistical notion of dimensionality reduction, through matrix algebraic methods, has been proposed as a means to access these synergies—of finding the lower functional dimensions amid the larger number of measured dimensions.

The idea of using dimensionality-reduction techniques to estimate synergies has been around for almost a half century. Easton (1972), described in Turvey (1977, 2007), sought to articulate the weaving of unit reflexes into more complex performances. Turvey identified this general approach with finding the "basis" over which the system is actually operating. The algebraic sense of basis is that we have a set of unit vectors (or other quantitative components) that preserve the original variability of the measured data but can be represented in fewer (perhaps many fewer) dimensions. The basis emerges from functional control of the reduction of degrees of freedom and can be interpreted both in a theoretical sense but also quantitatively in the form of specific statistical procedures that can extract this basis. For example, d'Avella, Saltiel, and Bizzi (2003) used matrix factorization methods to find low-dimensional representations of muscle synergies in frog-leg actions, such as swimming, jumping, and so on.

The definition of synergies in this case can be rendered statistically in rather direct terms. A synergy is a collection of systematic but time-varying muscle activations that unfold during an organized action. These activations and their modulations can be represented quantitatively as dynamically weighted vectors in a lower dimensional output from some matrix factorization method.

Recent work further suggests this lower dimensional organization may operate through a synergy of muscle components, specifically in the speech domain.

[2]At the risk of seeming theoretically naive, I avoid a discussion of terms with almost equivalent extensions but perhaps subtly different connotations, including concinnity, coordinative structure, and so on, and group the key concepts under this one designation.

Work by Ghazanfar and colleagues (reviewed in Ghazanfar & Takahashi, 2014) gives intriguing clues that the homologous behavior in primates and monkeys may be the lip smack, which exhibits similar temporal organization, with key timescales operating at about 3–7Hz. This frequency band appears in the analysis of "natural statistics" in speech production by Chandrasekaran, Trubanova, Stillittano, Caplier, and Ghazanfar (2009). They identified correlation structure in the facial and vocal aspects of both artificially and naturally produced speech through detailed temporal analysis of existing databases.

These authors drew bold conclusions from this and related work: "Taken together, they indicate that the speech signal is, to some degree, agnostic with regard to the precise modality in which it is perceived" and that aspects of speech "emerge via the interactions between the brain, body, and environment," that experience is multimodal, and so is the basis of speech (Chandrasekaran et al., 2009, p. 15). Similarly, Gick and colleagues recently argued that discovering these synergies in speech is a crucial next step in understanding how human vocal control works (Gick & Stavness, 2013). This recent work serves as further momentum to important prior work—also likely well known to readers—that synergistic properties of the speech system can be tested in a variety of elegant studies (e.g., Kelso, Tuller, Vatikiotis-Bateson, & Fowler, 1984).

AN INTEGRATIVE SYSTEM-LEVEL STRATEGY

The study of speech is a productive domain in which to explore synergies, as it intrinsically involves engagement of a suite of muscle combinations. I began this article, however, with a description by Enfield (2013) that sees contributions to natural discourse as composite wholes, which include speech and "much more besides." If it's true that coordinated linguistic performance is synergistic more broadly, how do we discern synergies at this broader linguistic level of analysis? Fusaroli, Rączaszek-Leonardi, and Tylén (2014) observed that "in order to move from simple motor control to dialog as a functional and inter-individually defined system, one is presented with the difficult task of specifying its function. Most often people engage in dialog with specific cooperative purposes" (p. 151). We need to bridge lower and higher scales of linguistic description, from phonetic control to dialog functions.[3]

An interesting comparison can be drawn in recent linguistic theory. For example, Jackendoff (2002) articulated the need to move beyond "syntactocentrism" and

[3]It does seem jarring to put it this way, bridging such vast scales. The point I am making is that humans weave phonetic control into discourse contexts, so why shouldn't our science of language try to do so, too?

identify how other aspects of language can be productive and, it is important to note, how they are interfaced in order to function together. One strategy to solve this problem is to add "interface processes" to account for the integrated system. It is encouraging that theoretical developments in linguistics are recognizing the multidimensional basis for natural language performance, and this strategy also takes a kind of integrative approach: it takes phonology and conceptual structure on either ends of a linguistic system, with syntax as one part of this broad spectrum of linguistic properties.

Ecological readers are probably less inclined to embrace a foundation of abstract linguistic structures out of the explanatory gate. Instead, we desire foundations from more directly measurable variables, such as at the level of sound systems (e.g., Port, 2007). But we can take a similar integrative strategy in a highly data-driven approach. The natural structure of human interaction involves discourse functions, or "oral genres" (Busch, 2007): giving directions, explaining a concept, describing an episodic memory. These unfold at a slower timescale. On the opposite end, we have the rapidly changing behaviors that compose these performances—the ebb and flow of body movement, the staccato on and off states of the vocal system, the distribution of eye movements, and so on. In between these two timescales—the slowest and the fastest—lies an array of processes that we typically associate with language, from the selection of words, their linearization into sentences, and so on.

A radically integrative approach would seek to tether the two rather disparate timescales described here.[4] We may focus on the slowest and fastest timescales as a beginning point for extracting systematic relationships between them and articulating the synergies that lie between. The methodological motivation is explicitly statistical and data driven. The various systems participating in linguistic performance interact, and as such, carry variance about each other and the rest of the system—even among relatively "incidental" behaviors, such as eye movements or subtle patterns of speech. As a result, there is some lower dimensional characterization of the system that can be extracted from these behaviors. These can be represented statistically through dimensionality-reduction techniques used similarly and productively all around cognitive science, such as Principal Component Analysis, Latent Semantic Analysis (Landauer, McNamara, Dennis, & Kintsch, 2013), Multidimensional Scaling

[4]Something I deliberately avoid here are the many examples in sociolinguistics and related domains where low-level variables have indeed been correlated in functional ways in discourse contexts—for example, the well-known discussion of Lindblom (1990) involving adaptive speech in different communicative contexts. I mean something quite different here, but with more space, it would be worth elaborating this relationship further. The integrative approach argued for here seeks to find system-level characterizations of these relationships expressed in terms of the nature of the underlying system.

(Hout, Papesh, & Goldinger, 2013), Independent Component Analysis (Delorme & Makeig, 2004), Singular Spectrum Analysis (Golyandina & Zhigljavsky, 2013), and so on. In other words, "synergic components" can be quantified by finding and characterizing the system's lower dimensional basis, from the top to the bottom.

AN OBSERVER-CENTERED APPROACH

Many new questions can be posed by quantitatively linking the dynamic structure of lower level behavioral variables to high-level discourse functions and exploring how these converge in an organized fashion in natural linguistic performances. We have begun a research project in this vein by starting to map eye movements, body movement, and vocal modulation onto distinct discourse modes: giving directions, explaining an abstract concept, providing narrative of a recent memory, and so on.

Before sharing a few more methodological details of this research strategy, which space restricts here, let me further justify this integrative approach from an ecological perspective. For theoretical simplicity, especially for an assuredly theoretical readership, I take a purely observer-centric approach to defining the research context. This is easy to do and avoids ontological problems of such terms of behavioral "channels" or "processes." An observer of a natural interaction can carry out a coarse-grained coding of what is taking place. For example, an observer may determine that one person is explaining to another, or offering directions to another, and so on.[5] This can be referred to as a "discourse measure." An observer may also make note of the particular words that are being used in the interaction, which is a measurement level "below" the first measurement just described—change is occurring at a faster timescale. At the fastest timescales, we may leverage automatic tools that almost obviate the observer altogether—tracking the eyes, vocal modulation, and so on.

This layering of linguistic measurement levels is an uncontested notion, and ecological researchers might embrace it if we make the following simplifying assumption: Take these only as candidate stable modes measured by an observer.[6] Articulating levels of analysis is a strategy that has been conducted in domains where synergies have been influential (e.g., Saltzman, 1979).

The quantitative definition in the previous section—theoretically crude but statistically productive—is that a synergy can be regarded as a set of time-varying

[5]This coarse-grained coding is scientifically viable because it can be found to be adequately intersubjectively stable across observers equated for particular backgrounds and so on.

[6]An elegant discussion in this same journal about "symbols" as explanatory "outposts" would also be worth considering here (Tabor, 2002).

lower dimensional functions that temporarily stabilize in the manifestation of a "higher level" unit of measure. For example, d'Avella et al. (2003) characterized distinct frog-leg functions, such as swimming or jumping, as a time-varying lower level set of physical muscle activations. They were able to use matrix factorization methods over the large-scale electromyographic measurements they took of the muscle groups and identify these time-varying functions. The extracted functions can even be used to mathematically model the organized actions.

It may be possible to do something like this in the linguistic case. When a human interaction partner engages in an organized bout of natural linguistic behavior—such as giving directions or explaining an abstract concept—it is manifested in a set of time-varying lower level behaviors that compose it. It is Enfield's (2013) "composite," and presumably, it exhibits systematic dynamic structure that marks it as a given discourse goal.

The work by d'Avella et al. (2003), and the complexity of the linguistic case, recommend a data-driven approach. As we describe briefly here, we have devised a multimodal data-recording system that can be used in a within-subject experimental design to track multiple low-level behavioral signatures during natural discourse in an unobtrusive way. Our goal is to extract multiple dynamic measures at this lower level—eyes, voice, body—and map their temporal structure onto discourse functions that we induce participants to engage in: explanation, narrative, directions, and so on.

FIRST STEPS IN THIS APPROACH

This data-driven approach is precisely the one we are taking in the lab to map low-level variables onto high-level ones. There is, of course, ongoing related work that is chasing similar goals, such as in multimodal development of artificial agents or analysis of social signaling (e.g., Brunet, Cowie, Heylen, Nijholt, & Schröder, 2012; Kopp, van Welbergen, Yaghoubzadeh, & Buschmeier, 2014). The "integrative strategy" proposed here is explicitly theoretical and perhaps shamelessly open ended: seeking structure between the extreme ends of measurement scales may initiate next steps for identifying linguistic synergies in a coherent framework, for understanding how the parts hang together in naturalistic linguistic performance, and to develop a computational and theoretical framework that can link across levels of measurement.

We have adapted automated multimodal instrumentation to measure eye movements, vocal energy, and body movements in small segments of natural discourse performance, such as giving directions or explaining an abstract concept (see Figure 1). We have found so far that (a) low-level behavioral variation can indeed predict high-level discourse at sometimes a very high level

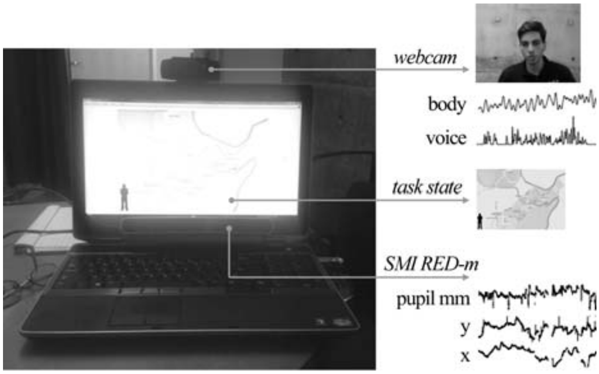

FIGURE 1 A representation of the research context. We utilize an eye tracker along with automated processing of webcam data to extract body, voice, and eye movement data while participants engage in instructed discourse tasks to a simulated task partner. Body movement is extracted using frame-differencing methods (Paxton & Dale, 2013). Vocal modulation is extracted by taking the average time-windowed amplitude of vocal energy. The Sensomotoric Instruments (SMI) RED-m eye tracker extracts a variety of measures at 120 Hz. These provide the low-level behavior in a search for a lower dimensional representation that can cluster or predict what discourse function the participants are engaged in. The task state is stable for each discourse instruction.

of accuracy, yet (b) participants create dynamic patterns on the fly, in the sense that one person's extracted synergies cannot easily predict the patterns of behavior in another participant (Dale, Paxton, & Gupta, 2015). Space limits the presentation of details of this ongoing work, but in early analyses we find promising predictive relationships between the lower and high levels and an intriguing amount of individual variability that suggests the manifestation of these synergies may be heavily task based.[7]

This approach, though still in a preliminary stage, may be useful for pursuing a system-level description of language performance—how the broader system is constituted by interactions among its parts (see Figure 2). Language is learned exclusively in a social context. It is learned with an expectation that we will be communicating with another person. Because of the richness of language in context, an approach by synthesis would seem important for explanatory progress

[7]This suggests intriguing relationships to task-driven factors as described by Saltzman and colleagues in influential work on motor control, where perhaps even linguistic performance can be seen as having intrinsic task-based dynamics rather than being generated purely by a complete endogenous cognitive plan, as "an invariant control structure that is specified dynamically according to task requirements and ... gives rise to diverse kinematic consequences" (Saltzman & Kelso, 1987, p. 4). The task context in this case would involve instructions, of course, but also an incremental context constructed by the participant herself or himself.

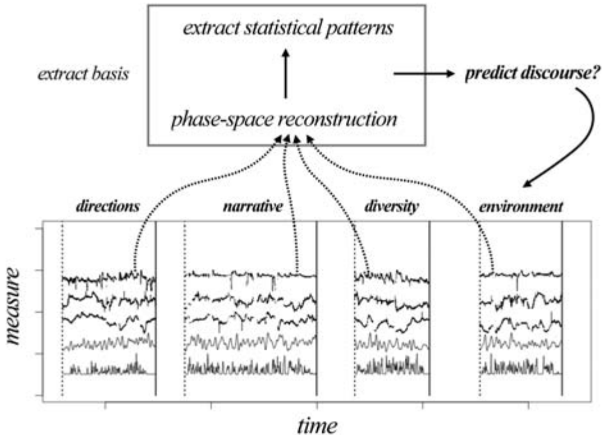

FIGURE 2 A caricature of the data-driven approach in our preliminary investigations. Our initial investigation utilizes phase-space reconstruction methods with a dimensionality-reduction technique known as singular spectrum analysis (SSA; Golyandina & Zhigljavsky, 2013). Participants are instructed to speak on a variety of topics, from giving directions (concrete topic) to describing the diversity on campus (abstract topic). We seek lower dimensional predictors, via SSA, to determine what dimensionality is capable of predicting the discourse modes.

and theoretical integration. I consider a few examples of potential empirical progress here.

- *Dimensionality of natural linguistic performance*. Intriguing work by Stephen, Dixon, and Isenhower (2009) suggests that cognitive performance can be studied even by extracting signatures from fast-changing behavioral measures, such as hand movement. In the dimensionality-reduction approach, something similar would be possible. We might pursue the dimensionality of soft-assembled control in various linguistic contexts and thereby pose new questions. When participants provide abstract descriptions, does this induce looser "cognitive control" in that the dimensionality inferred statistically is higher than in other discourse functions, such as giving directions? Are referent-coupled natural performances, such as giving directions, more statistically constrained, operating at a yet lower dimensionality?
- *How modalities are weaved*. Similarly, it is quite interesting to many language researchers to understand the role of eye movements, body movements such as gesture, and vocal control. Indeed, each of these levels invokes its own veritable literature of research on the topic. The development of statistical models, as proposed here, could yield new knowledge about the

role of these parts of natural performance but also the manner in which they are woven during distinct discourse functions.

- *Integrating the middle levels*. With more space, it could be argued that starting at the ends of the timescale spectrum may present unique opportunities for taking subsequent steps of figuring the "middle scales" into our understanding. Understanding the dimensionality and structure of fast-changing behavioral variables in the context of discourse functions offers strong empirical constraints for broadening our focus on the rest of the system. In other words, theoretical proposals for how linguistic processes, such as word choice, are related to natural performance will have to satisfy the empirical constraints above and below this level of measurement. The integrative approach may present new constraints to facilitate this theoretical development about natural performance.
- *Applied benefits*. The obvious applied benefits of this approach are that knowing something about subtle dynamic signatures—ones that are connoted by synergistic "control"—may be signatures that can be put to use in wearable computing systems that can track low-level signatures in pervasive high-level contexts (Paxton, Rodriguez, & Dale, 2015).
- *Monologue versus dialogue*. An obvious limitation to how we have begun this project is to focus on what is essentially monological: the performance of a single person in a simulated interactive context. This is a simplifying initial first step, as we believe the synergistic approach will be especially interesting in the interactive case (Shockley, Richardson, & Dale, 2009). The data-driven approach may help us understand how the levels of behavior of one interlocutor help constrain and shape those of another, dynamically in time. Recent work suggests that this coupling process is critical to interactive performance, from the neural levels and beyond (Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012).
- *Grounding dialogue*. As noted earlier, there has long been an interest in the multimodal and embodied nature of language, but I argue forcefully that it is beckoning for a more data-driven, quantitative approach. The results of a data-driven approach may permit a clearer understanding of how human dialogue—incredibly complex and abstract by many descriptions—is anchored statistically in its physical instantiation. This would not solve the grounding problem, which some have argued has already been solved (Steels, 2008), but it would drive new questions about how abstract levels of analysis in dialogue, still mostly inaccessible to automated quantification or data-driven analysis, may be usefully grounded through statistical means.

A data-driven approach that articulates statistical relationships among levels of measurement in language may bridge quite different forms of "representation." We are banking on it and conducting new studies described only briefly here.

We hope that the result is at least another step toward a broader understanding of the way that various aspects of natural language performance work together during that performance. Put simply, we hope to gain some insight into the structure of Enfield's (2013) composites.

REFERENCES

Bernstein, N. A. (1967). *The coordination and regulation of movements*. Oxford, UK: Pergamon Press.

Brunet, P. M., Cowie, R., Heylen, D., Nijholt, A., & Schröder, M. (2012). Conceptual frameworks for multimodal social signal processing. *Journal on Multimodal User Interfaces*, *6*, 95–99.

Busch, M. W. (2007). *Task-based pedagogical activities as oral genres: A systemic functional linguistic analysis* (Unpublished doctoral dissertation). Ontario Institute for Studies in Education, University of Toronto, Toronto, Canada.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*(7), e1000436.

Dale, R., Paxton, A., Gupta, S. (2015). *Predicting discourse from ensembles of non-verbal behavior.* Manuscript in preparation.

d'Avella, A., Saltiel, P., & Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nature Neuroscience*, *6*(3), 300–308.

Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Easton, T. A. (1972). On the normal use of reflexes: The hypothesis that reflexes form the basic language of the motor program permits simple, flexible specifications of voluntary movements and allows fruitful speculation. *American Scientist*, *60*, 591–599.

Enfield, N. J. (2013). *Relationship thinking: Agency, enchrony, and human sociality*. Oxford, UK: Oxford University Press.

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157.

Ghazanfar, A. A., & Takahashi, D. Y. (2014). Facial expressions and the evolution of the speech rhythm. *Journal of Cognitive Neuroscience*, *26*(6), 1196–1207.

Gick, B., & Stavness, I. (2013). Modularizing speech. *Frontiers in Psychology*, *4*, 977.

Golyandina, N., & Zhigljavsky, A. (2013). *Singular spectrum analysis for time series*. New York, NY: Springer Science & Business Media.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121.

Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2013). Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(1), 93–103.

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford, UK: Oxford University Press.

Kelso, J. S. (2009). Synergies: Atoms of brain and behavior. In D. Sternad (Ed.), *Progress in motor control* (pp. 83–91). Heidelberg, Germany: Springer.

Kelso, J. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(6), 812.

Kopp, S., van Welbergen, H., Yaghoubzadeh, R., & Buschmeier, H. (2014). An architecture for fluid real-time conversational agents: Integrating incremental output generation and input processing. *Journal on Multimodal User Interfaces*, *8*(1), 97–108.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2013). *Handbook of latent semantic analysis*. Mahwah, NJ: Psychology Press.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht, the Netherlands: Springer.

Paxton, A., & Dale, R. (2013). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*, *45*(2), 329–343.

Paxton, A., Rodriguez, K., & Dale, R. (2015). PsyGlass: Capitalizing on Google Glass for naturalistic data collection. *Behavior Research Methods*. Advance online publication. doi:10.3758/s13428-015-0586-z

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology*, *25*(2), 143–170.

Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science*, *33*(8), 1468–1482.

Richardson, M. J., Dale, R., & Marsh, K. (2014). Complex dynamical systems in social and personality psychology: Theory, modeling and analysis. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–282). New York, NY: Cambridge University Press.

Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in Psychology*, *2*, 38.

Saltzman, E. (1979). Levels of sensorimotor representation. *Journal of Mathematical Psychology*, *20*(2), 91–163.

Saltzman, E., & Kelso, J. (1987). Skilled actions: A task-dynamic approach. *Psychological Review*, *94*(1), 84.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*(2), 305–319.

Steels, L. (2008). The symbol grounding problem has been solved. So what's next? In M. D. Vega, G. Glenberg, & A. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 223–244). Oxford, UK: Oxford University Press.

Stephen, D. G., Dixon, J. A., & Isenhower, R. W. (2009). Dynamics of representational change: Entropy, action, and cognition. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1811.

Tabor, W. (2002). The value of symbolic computation. *Ecological Psychology*, *14*, 21–51.

Turvey, M. T (1977). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting and knowing: Toward an ecological physiology* (pp. 211–265). Mahwah, NJ: Erlbaum.

Turvey, M. T. (2007). Action and perception at the level of synergies. *Human Movement Science*, *26*(4), 657–697.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, *132*(3), 331–350.