This is a contribution from *Experience, Variation and Generalization. Learning a first language*.
Edited by Inbal Arnon and Eve V. Clark.

Genetic variation and individual differences in language

Jennifer B. Misyak & Morten H. Christiansen
Cornell University/Cornell University & Santa Fe Institute

Drawing upon recent conceptual and analytical trends in molecular genetics, this chapter explores how molecular genetics might be used to elucidate language differences in normally developing individuals, including variation in linguistic generalization skills. We propose an endophenotype approach, complemented with learning-based tasks and process measures, as a step forward in studying variation in the underlying brain mechanisms that support normal language learning and processing. We discuss how well-formulated and theoretically grounded endophenotypes may offer greater conceptual clarity, objectivity, and precision over many conventional assessments used in language genetics. This framework aims to uncover the biological part of an emergentist account to language, and provides a complement to investigations into more experientially based factors.

Keywords: Endophenotype; genetic variation; individual differences

1. Introduction

Individual differences in language are ubiquitous. From native language development (e.g. Bates, Dale & Thal 1995) to secondary language learning (e.g. Dörnyei 2005), and across children's and adults' language processing (e.g. Farmer, Misyak & Christiansen, in press; MacDonald & Christiansen 2002), substantial and systematic variation is to be expected. Such variation reflects the interdependent influences of environmental and genetic sources across individuals. Complementing other approaches in this volume that consider experiential factors, we focus in this chapter on how molecular genetics might be utilized to widen our understanding of individual differences in language skill. While assuming an implicit focus on variation in oral language (where applicable), our main arguments and points are also relevant to more specific attempts to understand differences in linguistic generalization.

Our emphasis throughout is on an endophenotype approach with learning-based tasks, as a way of operationalizing an 'emergentist genetics' that may

elucidate language variation in normally developing individuals (including in linguistic generalization). Endophenotypes are quantitative measurements of intermediate *phenotypes* (behavioral traits) that mediate between *genetic antecedents* (genes, gene products, and genetic factors) and more distal, complex phenotypes. We believe that the combination of endophenotypes and learning-based tasks (especially tasks with 'process' measures that tap into ongoing learning) represents a promising research strategy that avoids the limitations in many conventional approaches to genetic studies of language.

To set the groundwork for this strategy, we first identify some insights into language development afforded from behavioral genetics findings. We then clarify how an emergentist genetics, within which our approach is situated, fundamentally contrasts with linguistic-nativist construals of the genetics of language. With these preliminaries in place, we then identify how modern conceptual trends in molecular genetics appear well aligned with emergentism. We describe the ways in which these modern trends translate into some interesting findings across normal language and cognition. Finally, we argue that conventional phenotypic assessments of human language behaviors are incomplete and mostly underpowered for tracing the multifactorial and polygenic underpinnings (i.e. one trait may be affected by many genes) of complex traits such as language. Instead, we propose an endophenotype approach as a step towards remedying some existing shortcomings, and provide pointers to the nature and types of tasks (including tasks involving linguistic generalizations) suited for this approach.

1.1 Behavioral genetics and aspects of language development

Although the focus of this chapter is on the potential contribution of molecular genetics to the understanding of language development, research on behavioral genetics provides a useful starting point. Of key importance is that behavioral geneticists have estimated heritability coefficients for verbal ability to range from .37 to .82 (Bouchard & McGue 2003). More specifically, substantial heritability estimates obtain across diverse measures of language processing, including grammar, vocabulary, orthographic coding, phoneme awareness, phonological decoding, verbal memory, and word recognition (Gayán & Olson 2001; Kovas et al. 2005). Crucially, however, additional findings indicate that early syntactic and lexical development are affected by largely the same genetic factors and vary together (Dale et al. 2000; Dionne et al. 2003). Much of the variance for syntactic and lexical development is influenced by shared environmental factors, suggesting a less prominent role for genetics during early language development. Additionally, the poorer language performance associated with specific

language impairment (SLI) appears to be primarily influenced by environmental sources among those children with SLI in the absence of speech deficits (Bishop & Hayiou-Thomas 2008). Conversely, speech and phonological measures may vary separately from those of vocabulary-grammar, and may be more strongly influenced by genetic sources in early development (Bishop & Hayiou-Thomas 2008; Gayán & Olson 2001; Hayiou-Thomas et al. 2006).

In brief, meta-analysis from twin studies identifies the proportion of variance for language skills due to hereditary factors at .48 (Bouchard & McGue 2003; see also Stromswold 2001), implicating both a moderate environmental *and* hereditary contribution for linguistic proficiency. There may be some important differences in how these effects manifest themselves across different aspects of language, with some areas showing greater genetic influence across development than others. Nonetheless, as Bates et al. (1995: 97) noted, "At the risk of inviting accusations of radical centrism, we conclude that the variations observed in early language development are so large that they require substantial contributions from *both* genetic and environmental factors, with special emphasis on their interaction."

1.2 'The Nature of Nature': Emergentism

How should we view the interaction between genetic and environmental factors in language development? Dialogue in the language sciences has traditionally highlighted oppositions between linguistic-nativist or modularist perspectives, on the one hand, and those of emergentist, connectionist, or neo/neuro-constructivist views on the other. Within the former group, syntactic ability is thought to be subserved by a specialized neural substrate that develops universally across individuals, given appropriate (and relatively minimal) environmental inputs. This kind of innate endowment is exemplified in Chomsky's (1965, 1981) proposal of a Universal Grammar (UG), the formalization of a set of universal language-specific constraints. These views naturally accommodate proposals for mental modules (among which the language faculty is often prominently included) as dedicated, predetermined, and informationally-encapsulated systems to handle domain-specific inputs and cognitive representations (e.g. Atran 1998; Fodor 1983).

In contrast, emergentist, connectionist, neo/neuro-constructivist and similar views (e.g. Elman et al. 1996; Mareschal et al. 2007; Tomasello 2003) emphasize experiential processes that interact with the individual's general learning mechanisms over developmental time (as well as over evolutionary time: Christiansen & Chater 2008). In this way, the complex, species-typical patterns of behavior recognized as language can arise without entailing language-specific constraints or predetermined, domain-specific, neurobiological circuits. While such positions are

occasionally miscast as 'tabula rasa' empiricism, there is no theoretical requisite for eschewing some genetic biases. Amid an interactive developmental and ecological backdrop, such theories can assign a powerful role to small initial biases of the learning system. In principle, these biases can be architectural or chronotropic in form (Elman et al. 1996), and reside along a continuum of varying degrees of computational (connectionist) nativism (Clark 1993).

These two classes of views differ in their characterization of internal knowledge (Clark & Misyak 2009), with implications for the level, manner, and specificity with which genetic effects may operate. Essentially, linguistic-nativist theories suggest a strong structurally and/or functionally specific biological basis for language. Conversely, the emergentist perspective points to small, germinal biases or broad intrinsic constraints influencing more general, 'low-level' biological systems of relevance to language. These are thus diverse proposals for the ways in which language may have genetic foundations and influences.

Although linguistic-nativism has traditionally appropriated genetics within the purview of 'Nature,' we side with the view that genes more realistically have explanatory relevance as biases embedded within a normative biochemical and ecological context (Clark & Misyak 2009). While investigating these interdependent influences in language requires a multifaceted, broad approach, this chapter examines a small (but vital) subspace in such an account. In delving next into the molecular genetics literature, we should not be seen as privileging the molecular level (nor genetic contributions), but rather as exploring the potential for unraveling the *biological part* of an emergentist account that will require synergistic interconnections with other investigations of environmental factors (as discussed elsewhere in this volume).

2. Genes and behavior

Although we will not survey the history of molecular genetics, it is useful to highlight some constants and shifts in current thinking, which actually bring modern genetics into better accord with an emergentist framework. It has been recognized for some time that particular genotypes and particular environments co-occur with one another at appreciable frequencies (see further Plomin, DeFries & Loehlin 1977; for an example in language development, see Gilger et al. 2001). A practical consequence of these gene-environment correlations is an increase in phenotypic variance. Conceptually, this indicates the presence of amplifying or compensatory couplings, and underscores the complexity of disentangling putative 'environmental' and 'genetic' influences. Some researchers consider such correlations as

part of the normative manner in which genes underlying polygenic human traits are expressed, i.e. genes typically work by recruiting particular environments.

A step further in interactivity, *epistasis* (or, gene-gene interaction, a term introduced in the twentieth century; cf. Cordell 2002) continues to receive attention. In its simplest sense, 'compositional epistasis' refers to the action of a dominant allele at one locus (gene) masking the effect of other allele(s) at different loci (genes). More complexly, this term can also be extended to refer to non-additive influences for quantitative phenotypes (i.e. 'statistical epistasis'). Beyond intra-gene relations, genotype-environment interactions (*GxE*) in the statistical analysis-of-variance sense have also traditionally been acknowledged (see Dempfle et al. 2008 for a review) and have lately yielded surprising findings in cognitive development (e.g. Caspi et al. 2007), though not yet in language.

Tracing schools of thought in evolutionary biology, Gontier (2008) suggests we are in the midst of a paradigm shift from a traditional view of genes as inherently encoding and transmitting information about morphological form (and function) to a newer conceptualization in which information is viewed as an "exherent, emerging property of genes." Gontier credits the roots of this "information-emerging metaphor" to systems theory and Schrödinger's construal of genes as underlying "difference in properties," rather than acting as self-contained coding-entities (Schrödinger 2000, as cited in Gontier 2008: 177; see also Wheeler & Clark 1999).

Gontier's identification of this paradigm shift merges with the transformative modern notion of epigenesis that is developing within molecular genetics. Epigenesis frequently refers to heritable cell changes (such as modifications of chromatin or DNA methylation) independent of any corresponding change in the DNA sequence. It broadly subsumes the differential expression of genes in tissue and the regulation of genes over time (Carey 2003).

By incorporating temporal dynamics into the picture, epigenetics conceptually accommodates the *emergent aspects* of *gene-environment interactions* at the biological level (i.e. the integration of changes in the environment and the expressed genotype over developmental time). This is consistent with the recognition that much of DNA fulfills regulatory purposes rather than structural functions (Gottesman & Hanson 2005). In addition, these epigenetic processes are primarily probabilistic (e.g. Gottlieb 1998), rather than strictly predetermined, and substantially informed by internal/external environmental signals. Thus there is a confluence of genetic, environmental, stochastic, and epigenetic factors in the current understanding of the genotype-phenotype relationship. While this introduces added complexity – the complexity of co-determining interactions among coupled organisms and environments, as well as, technically, decouplings between genetic (DNA sequencing) variations and epigenetic systems (see also Jablonka & Lamb 2002) – it appears a

necessary conceptualization that accords with modern data. It also resonates with the nature of adaptive, dynamic biological systems and their activity-dependent modification. Equally clearly, it does not pre-empt the role of genetic constraints or their investigation.

Dovetailing with the 'information-emerging' metaphor and epigenetics are newer analytical techniques and models, including the convergence of quantitative and molecular genetics in QTL (quantitative trait loci) research. Under a QTL model (Plomin, Owen & McGuffin 1994), complex traits are assumed to be associated with many genes, such that the varying effect size for any single gene will likely be small. 'QTLs' refer to such genes, whose integrative genetic effects represent probabilistic propensities. As an example of the ramifications of 'probabilistic,' among the strongest and well-established of complex behavioral associations to date, variation in the type 4 allele of the apolipoprotein E (APOEε4) with Alzheimer's disease is neither necessary nor sufficient to produce the behavioral outcome of late-onset dementia (Plomin et al. 2008). Therefore, genes (loci) comprising QTLs associated with variation in a quantitative cognitive-behavioral trait may be construed as probabilistic (not predetermined) and interchangeable contributions for informing brain development.

An important implication of the QTL model's premise that many genes affect a behavior (*'polygenicity'*; see further Kovas & Plomin 2006) is that there will exist a continuous, quantitative distribution for the phenotype. To make this clear, we adapt an example from Plomin et al. (2008): Suppose a single gene with two potential, equal-frequency alleles of cumulative influence. This allows for three possible genotypes, with a distribution among the population of .25, .50, and .25 (i.e. 25% homozygous for allele 1, 25% homozygous for allele 2, 50% heterozygous). Assuming polygenicity (many genes influencing behavior), we can extend the example to two genes, which results in nine genotype-pairings and five phenotypes with a distribution of 1/9, 2/9, 3/9, 2/9, 1/9. At three genes (twenty-seven genotypes, seven phenotypes), this approaches a normal distribution, even without considering non-equal alleles and environmental variation.

3. Molecular genetic studies in language and cognition

Conceptual shifts in molecular genetics have been accompanied so far by a number of interesting findings. First, language-related studies often have implications for normative language skills, consistent with the notion that genetic influences supporting language abilities may be continuous with those involved in language disabilities at the lower end of the normal distribution (Tomblin & Christiansen 2009). This follows from the QTL model's premise of polygenicity (see above),

though it precludes rare neuro-developmental disorders with very low prevalence (≤.01% per QTL criteria; Butcher, Kennedy & Plomin 2006). This is compatible with other general trends that regard many putative disorders and exceptional manifestations of skills as situated at the extremes of a quantitative continuum. This thinking is aptly captured by the aphorism, 'the abnormal is normal' (Plomin & Kovas 2005).

Second, as Gontier (2008) observed, it is becoming apparent that major genes previously implicated in language and cognition (e.g. *FOXP2, ASPM, MCPH1*) do not unequivocally code for morphological features of trait-specific behavior as classically expected. Many such genes are regulatory and '*pleiotropic*' (i.e. a single gene may affect many different traits). Thus, they often have diffuse and graded effects on several neurobiological systems underlying behavior.

Furthermore, recent studies connecting genetic effects to normative linguistic performance generally point towards sources of common allelic variation and to nonobvious influences, rather than to abnormal genetic mutations and dramatic effects. For instance, mutation in primary microcephaly-associated genes such as *ASPM* (abnormal spindle-like microcephaly-associated) and *Microcephalin* (*MCPH1)* leads to a rare neurodevelopmental disorder producing a reduced-size brain with severe mental deficits (Bond & Woods 2006). But it is unclear whether there is any association with brain size variation in the normal population (e.g. Rimol et al. 2010; Woods et al. 2006). Nonetheless, derived alleles of *ASPM* and *Microcephalin* are strongly associated with languages employing a single-tier phonological system (rather than two-tiers, with tonal contrasts) (Dediu & Ladd 2007). Along with phenotypic association of *ASPM* to oral reading comprehension and phonology measures of second-grade children, these findings suggest a subtle role for common allelic variations in contributing to language development (Christiansen, Kelsey & Tomblin 2008), in the form of differences in phonological sequential learning biases.

Finally, a caveat: only a few molecular studies of normal language abilities have been conducted and with limited results. This reflects some of the methodological limitations in using conventional phenotype assessments, along with QTL implications of small effect sizes. However, future genetic studies of language could likely benefit from adopting an endophenotype strategy.

4. Towards an endophenotype approach

The attempt to trace polygenetic antecedents that contribute to normal variation in language (and other cognitive behaviors) may appear rather daunting. Indeed, it contrasts starkly with the success of traditional linkage studies reporting

single-gene discoveries for simple (albeit rare) Mendelian diseases. With complex phenotypes as the norm rather than exception, one insight into their biological underpinnings bypasses direct attempts at genetic association with the phenotypes themselves (i.e. the skipping of multiple levels from gene or gene product to high-level cognitive behavior) and derives its inspiration from insect biology.

John and Lewis (1966) introduced an 'endophenotype' (*endo* meaning "inside;" *pheno* meaning "show") as a neglected component for explaining variance in geographical distributions among insect populations, contrasting it with macroscopic or *exophenotypic* influences that other researchers attended to at the time. They urged a fuller examination of "not the exophenotype but the endophenotype, not the obvious and external but the microscopic and internal" (1966: 720). So when Gottesman and Shields (1973) adapted this term, they underscored a measurable *internal phenotype* that could be discovered by "biochemical test or microscopic examination." It was hoped that such tractable measurements would aid in the genetics of seemingly intractable-to-study psychiatric disorders such as schizophrenia. Within the last few years, the endophenotype concept has commanded greater attention, discussion, and usage. However, the notion of endophenotypes has yet to be typically extended beyond clinical applications to normal behavior as the primary scientific focus.

4.1 Concept

An endophenotype, or internal phenotype, is intended to mediate between an external, complex manifestation of behavior and its etiologically complex genetic antecedents. Endophenotypes are thus quantitative traits 'not obvious to the unaided eye' that are presumed to lie anywhere 'intermediate' along the causal pathways between genes (or gene products) and phenotypic traits. As endophenotypes are by definition putatively closer to genes, it is hoped that they will facilitate the discovery of genetic variants causal to behavioral traits, and, in the case of psychiatric or medical applications, refine or stratify diagnostic classification for poorly defined heterogeneous disorders using neurobiological information. Classical examples in the medical literature, where biometric assessments are prominent, are serum cholesterol-level assays as a predictor of cardiovascular disease or glucose tolerance tests as indicators of diabetes type I susceptibility; both have heritable contributions and are associated with genetic loci (Topol et al. 2006; Tuomilehto-Wolf et al. 1993).

An essential feature of endophenotypes is that they relate to quantitative and continuous (not categorical) variance. This is true regardless of phenotypic 'expression' (i.e. they have some degree of stability or state-independence). They should lend greater objective precision to genetic investigations than assessments based on

an externally diffuse behavior. And along with more tractability, they should yield genetic associations with larger effect sizes than attempts at more distal mappings.

Another feature of endophenotypes is the diversity of 'levels' at which they are characterized. Given the multicausal and multilevel pathways from molecules to traits, endophenotypes can be quantitative measurements residing at any of the intervening levels: e.g. (neuro)physiological, biochemical, endocrinological, (neuro)anatomical, and (neuro)cognitive. Hence, the synonymous use in the literature of endophenotypes as 'internal phenotypes' or 'intermediate phenotypes' indexes their internal status and medial position, respectively.

4.2 Merits

When properly implemented, endophenotypes would appear to possess strong advantages as a research tool. They have been theorized to confer greater statistical power and tractability to genetic designs. Since they are closer to genetic antecedents than the phenotype in question, endophenotypes promise to involve correspondingly fewer loci and/or greater effect size for successful association. This assumption of power rests on genetic proximity, not simplicity of biological level. For instance, variation in gene expression profiles, at the level of transcriptome (transcribed RNA) analysis in functional genomics, is *very* basic, but only modestly (and polygenicly) heritable (see Plomin et al. 2008).

Another merit of endophenotypes, particularly brain-based endophenotypes, is that their intermediary level status can provide converging information about underlying neural systems that support a cognitive behavior. Furthermore, as quantitative assessments rather than as qualitative (and frequently subjective) measurements, they have the added advantage of objectivity. Given the substantial influence of environment and compensatory factors on behavior, endophenotypes are potentially 'cleaner' in their identification with underlying learning-based neural mechanisms and related genetic variants. For psychopathology research, endophenotypes may also possess earlier temporal sensitivity in diagnosing specific conditions. In sum, an endophenotype strategy – with greater conceptual clarity, objectivity, and precision – would appear indispensable for molecular studies of normal language.

5. Possible language endophenotypes

Endophenotypes in their short history have met with some success in the psychiatric genetics field where they originated (Burmeister, McInnis & Zollner 2008). They have been employed in the discovery of novel genes and in

revealing functional consequences of susceptibility alleles on (neuro)cognitive measures (Walters & Owen 2007). For instance, the identification of two new genes (*GABRA2* and *CHRM2*) associated with a predisposition to alcohol dependence was attributable to the use of electrophysiological endophenotype measures, which enabled finer localization of the linkage signal and greater statistical power (Dick et al. 2006). Some useful outgrowths of the endophenotype approach are that: (1) endophenotypes have been associated with quantified symptomatology rather than discrete diagnostic category labels, and (2) the approach has revealed common variation within the normal population, not just within the clinical sub-group.

However, few endophenotypes have been employed in studies aimed at elucidating underlying neural mechanisms and genetic sources of variance for normal language abilities. Some exceptions, though, include behavioral genetics work on the heritability of nonword repetition performance (Bishop 2009). Such work supports the possibility that phonological short-term memory (PSTM), measured by a nonword repetition task, may serve as an endophenotype to distinguish unimpaired language performance from common manifestations of impaired language performance (i.e. typical forms of SLI). Vernes et al. (2008) also used this endophenotype of PSTM for identifying and associating single-nucleotide polymorphisms (SNPs) in *CNTNAP2* (a gene down-regulated by *FOXP2*) with common forms of language impairment. The success of PSTM as an endophenotype for understanding genetic factors relevant to the full spectrum of language abilities has yet to be fully investigated. We envision, though, that other endophenotype candidates, such as those related to general sequential processing mechanisms (which may also play a role in supporting linguistic generalizations; e.g. Christiansen & Chater 2008) and with some of the properties described below, may hold greater utility in accounting for a large amount of normal language variation.

5.1 Neuroimaging techniques

One level at which endophenotypes can enter into the picture is through contact with cognitive neuroscience. In fact, 'imaging genomics' (as it has been dubbed) appears to be rapidly growing, and the use of neurocognitive endophenotypes generally partners well with and enhances sensitivity of cognitive-behavioral endophenotypes/assays (Goldberg & Weinberger 2004).

With structural imaging techniques, magnetic resonance imaging (MRI) markers of brain structure have been established as heritable. Very high heritability for gray matter distribution in Broca's and Wernicke's areas has also been reported (see Toga & Thompson 2005). For other anatomical endophenotypes, more local measures could be employed. For instance, specific gray matter (and generic white

matter) has been presented as an endophenotype for schizophrenia and bipolar disorder using voxel mapping (McDonald et al. 2004).

However, a systems-level account may accord better with our appreciation of the brain's interactive organization, and some researchers have targeted endophenotypes at this level of description. Accordingly, MRI, functional magnetic resonance imaging (fMRI), and event-related potentials (ERPs) have been fairly widely used to ascertain neural responses to cognitive tasks in an effort to bridge brain-level variability to underlying genetic variation, and to relate brain activity to cognitive behavior. In some cases a relationship between genetic and brain-based variation may manifest itself in the absence of any relation between neural and cognitive variation owing to compensatory factors. Brain-based endophenotypes can therefore serve as an additionally informative entry-point for understanding causal links that might otherwise remain unobserved (see Green et al. 2008).

5.2 Process measures and learning-based tasks

Whether neuro-cognitive or cognitive-behavioral endophenotypes are proposed, 'process' measures (assessments that tap into ongoing cognitive processes) are better poised to contribute to understanding genetic variants associated with brain-systems supporting language than 'product'-oriented measures (i.e. assessments of static outcomes of learning/processing). Process-based measures lend more sensitivity, ecological validity, and objectivity to study designs than product-oriented evaluations. *Learning-based tasks* for endophenotypes are also better controls for environmental variations than many traditional tasks employed in language research. That is, experimental learning tasks using carefully designed novel stimuli promise to do better at tapping into key processing skills that are likely to be less influenced by participants' previous environmental exposures. Learning-based tasks and process measures complement each other (and can be also instantiated through, or used in conjunction with, systems-level neuroimaging techniques). Specific to exploring differences in linguistic generalization, we envision that learning-based tasks could be designed or adapted to measure well-formulated endophenotypes supported by the empirical literature.

The adoption of process measures and learning-based tasks represents a significant departure from conventional assessments employed in genetic studies of (ab-)normal language. For instance, previous studies of language in behavioral genetics have relied on school/parental reports, questionnaires, telephone interviews, or standardized and non-standardized test measures (Plomin & Kovas 2005; Stromswold 2001). These assessments tend to lack dimensionality and can be highly subjective. As such, they may contribute considerable (non-genetic) variance that masks genetic effects on underlying neurocognitive systems.

Product-oriented assessments also represent the behavioral product of cumulative learning, informed by both environmental and genetic influences. Many were specifically devised as 'achievement' measures, thus possibly conflating or masking potential genetic influences with patterns of environmental variance in producing learning outcomes. These conventional measures rarely, if at all, allow for the study of more intricate aspects of language, such as linguistic generalization.

5.3 Incorporating developmental change and other challenges

Given the dynamic interplay among genetic, environmental, stochastic, and epigenetic factors, one would like to ask, 'How might the endophenotype change over time?' And more generally, 'What are the mechanisms by which genetic influences have their effect?' The short answer is that we need cross-sectional and longitudinal designs. But beyond this, in light of the novelty of the approach and the paucity of empirical data, we have no specific guidelines to offer at this time.

Developmental considerations and timing are clearly important, but it is difficult to know *a priori* exactly what to expect. While genetic influences become amplified over time (Plomin et al. 2008), it is also the case that cognitive profiles evidenced at a given developmental time-point cannot necessarily be used to infer cognitive profiles and their associated modulation by genetic factors at other points in time (Scerif & Karmiloff-Smith 2005). Furthermore, across the lifespan there are neural constraints in early development, and degenerative processes in very late adulthood. While early 'endpoints' are frequently given due consideration in studies, later 'endpoints' may be equally informative. For example, genetic factors relating to neuronal support and repair may manifest themselves most strongly as effects on cognitive performance in older age, such as may be seen in the influence of the apoE genotype and ER-alpha gene in older (but not young) populations (Greenwood & Parasuraman 2003). Developmental windows, timing of assessments on a micro-scale, and different allelic contexts likely play into the picture as well, and so underscore the need for both macro and micro developmental designs to inform future steps and flesh out a thoroughly emergent, dynamic perspective.

Lastly, we note that suitable endophenotypes need to be based on valid theoretical constructs, sensitive to individual differences, psychometrically strong, and employed in studies with sufficiently large samples. Failures to meet these criteria offer pitfalls for any task intended to assess an endophenotype. In particular, current neuroimaging tasks are generally cost-prohibitive, and current process measures sometimes have poor psychometric properties. (Product measures, too, can be lacking in these criteria, with potentially weaker ecological validity and theoretical grounding). We present these closing caveats not as barriers, but to promote caution and encourage researchers to tackle the many challenges remaining in the future.

6. Conclusion

In this chapter, we devoted our attention to a small problem-subspace of what is clearly a much larger and more arduous enterprise in elucidating variation in underlying brain mechanisms that support normal language abilities, including linguistic generalizations. Given newer conceptual and analytical approaches in molecular genetics, we have proposed an endophenotype strategy as a way to move forward, beyond the limitations in many traditional assessments in the language literature. Well-selected endophenotypes, with learning-based tasks and process measures, could lend greater power, tractability, conceptual clarity, measurement sensitivity, and objectivity to genetic designs aimed at illuminating underlying brain-based mechanisms for language behaviors. We conclude then with an unfinished, developmental picture and a provisional sense of our footing in this modern epigenetic landscape, simultaneously humbled and emboldened by the challenge of uncovering the role of genetic variation in accounting for individual differences in language.

Acknowledgments

We would like to thank Stewart McCauley, the editors, and two anonymous reviewers for their comments on a previous version of this chapter.

References

Atran, S. 1998. Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences* 21: 547–609.

Bates, E., Dale, P.S. & Thal, D. 1995. Individual differences and their implications for theories of language development. In *Handbook of Child Language,* P. Fletcher & B. MacWhinney (eds), 96–151. Oxford: Blackwell.

Bishop, D.V.M. 2009. What can developmental language impairment tell us about the genetic bases of syntax? In *Biological Foundations and Origin of Syntax,* D. Bickerton & E. Szathmáry (eds), 185–206. Cambridge MA: The MIT Press.

Bishop, D.V.M. & Hayiou-Thomas, M.E. 2008. Heritability of specific language impairment depends on diagnostic criteria. *Genes, Brain and Behavior* 7: 365–372.

Bond, J. & Woods, C.G. 2006. Cytoskeletal genes regulating brain size. *Current Opinion in Cell Biology* 18: 95–101.

Bouchard, T.J. & McGue, M. 2003. Genetic and environmental influences on human psychological differences. *Journal of Neurobiology* 54: 4–45.

Burmeister, M., McInnis, M.G. & Zollner, S. 2008. Psychiatric genetics: Progress amid controversy. *Nature Reviews Genetics* 9: 527–540.

Butcher, L.M., Kennedy, J.K.J. & Plomin, R. 2006. Generalist genes and cognitive neuroscience. *Current Opinion in Neurobiology* 16, 145–151.

Carey, G. 2003. *Human Genetics for the Social Sciences*. Thousand Oaks CA: Sage.
Caspi, A., Williams, B., Kim-Cohen, J., Craig, I.W., Milne, B.J., Poulton, R. et al. 2007. Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. *Proceedings of the National Academy of Sciences* 104: 18860–18865.
Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge MA: The MIT Press.
Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Christiansen, M.H. & Chater, N. 2008. Language as shaped by the brain. *Behavioral & Brain Sciences* 31: 489–558.
Christiansen, M.H., Kelsey, K. & Tomblin, J.B. 2008. Associations of ASPM genetic markers with language and non-verbal intelligence. Paper presented at the Boston University Conference on Language Development, Boston MA.
Clark, A. 1993. Minimal rationalism. *Mind* 102: 587–610.
Clark, A. & Misyak, J.B. 2009. Language, innateness and universals. In *Language Universals,* M.H. Christiansen, C. Collins & S. Edelman (eds), 253–260. Oxford: OUP.
Cordell, H.J. 2002. Epistasis: What it means, what it doesn't mean, and statistical methods to detect it in humans. *Human Molecular Genetics* 11: 2463–2468.
Dale, P.S., Dionne, G., Eley, T.C. & Plomin, R. 2000. Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language* 27: 619–642.
Dediu, D. & Ladd, D.R. 2007. Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences* 104: 10944–10949.
Dempfle, A., Scherag, A., Hein, R., Beckmann, L., Chang-Claude, J. & Schafer, H. 2008. Gene-environment interactions for complex traits: Definitions, methodological requirements and challenges. *European Journal of Human Genetics* 16: 1164–1172.
Dick, D., Jones, K., Saccone, N., Hinrichs, A., Wang, J., Goate, A. et al. 2006. Endophenotypes successfully lead to gene identification: Results from the collaborative study on the genetics of alcoholism. *Behavior Genetics* 36: 112–126.
Dionne, G., Dale, P.S., Boivin, M. & Plomin, R. 2003. Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development* 74: 394–412.
Dörnyei, Z. 2005. *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah NJ: Lawrence Erlbaum Associates.
Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge MA: The MIT Press.
Farmer, T.A., Misyak, J.B. & Christiansen, M.H. In press. Individual differences in sentence processing. In *Cambridge Handbook of Psycholinguistics*, M. Spivey, M. Joannisse & K. McRae (eds). Cambridge: CUP.
Fodor, J. 1983. *The Modularity of Mind*. Cambridge MA: The MIT Press.
Gayán, J. & Olson, R.K. 2001. Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. *Developmental Neuropsychology* 20: 483–507.
Gilger, J.W., Hsiu-Zu, H., Whipple, A.D. & Spitz, R. 2001. Genotype-environment correlations for language-related abilities: Implication for typical and atypical learners. *Journal of Learning Disabilities* 34: 492.
Goldberg, T.E. & Weinberger, D.R. 2004. Genes and the parsing of cognitive processes. *Trends in Cognitive Sciences* 8: 325–335.

Gontier, N. 2008. Genes, brains, and language: An epistemological examination of how genes can underlie human cognitive behavior. *Review of General Psychology* 12: 170–180.

Gottesman, I.I. & Hanson, D.R. 2005. Human development: Biological and genetic processes. *Annual Review of Psychology* 56: 263–286.

Gottesman, I.I. & Shields, J. 1973. Genetic theorizing and schizophrenia. *British Journal of Psychiatry* 122: 15–30.

Gottlieb, G. 1998. Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesis. *Psychological Review* 105: 792–802.

Green, A.E., Munafo, M.R., DeYoung, C.G., Fossella, J.A., Fan, J. & Gray, J.R. 2008. Using genetic data in cognitive neuroscience: From growing pains to genuine insights. *Nature Reviews Neuroscience* 9: 710–720.

Greenwood, P.M. & Parasuraman, R. 2003. Normal genetic variation, cognition, and aging. *Behavioral and Cognitive Neuroscience Reviews* 2: 278–306.

Hayiou-Thomas, M.E., Kovas, Y., Harlaar, N., Plomin, R., Bishop, D.V.M. & Dale, P.S. 2006. Common aetiology for diverse language skills in 4 1/2-year-old twins. *Journal of Child Language* 33: 339–368.

Jablonka, E. & Lamb, M.J. 2002. The changing concept of epigenetics. *Annals of the New York Academy of Sciences* 981: 82–96.

John, B. & Lewis, K.R. 1966. Chromosome variability and geographic distribution in insects. *Science* 152: 711–721.

Kovas, Y., Hayiou-Thomas, M.E., Oliver, B., Dale, P.S., Bishop, D.V.M. & Plomin, R. 2005. Genetic influences in different aspects of language development: The etiology of language skills in 4.5-year-old twins. *Child Development* 76: 632–651.

Kovas, Y. & Plomin, R. 2006. Generalist genes: Implications for the cognitive sciences. *Trends in Cognitive Sciences* 10: 198–203.

MacDonald, M.C. & Christiansen, M.H. 2002. Reassessing working memory: A comment on Just & Carpenter (1992) and Waters & Caplan (1996). *Psychological Review* 109: 35–54.

Mareschal, D., Johnson, M.H., Sirois, S., Spratling, M.W., Thomas, M.S.C. & Westermann, G. 2007. *Neuroconstructivism: How the Brain Constructs Cognition*. Oxford: OUP.

McDonald, C., Bullmore, E.T., Sham, P., Chitnis, X., Wickham, H., Bramon, E. et al. 2004. Association of genetic risks for schizophrenia and bipolar disorder with specific and generic brain structural endophenotypes. *Archives of General Psychiatry* 61: 974–984.

Plomin, R., DeFries, J.C., & Loehlin, J.C. 1977. Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin* 84: 309–322.

Plomin, R., DeFries, J.C., McClearn, G.E. & McGuffin, P. 2008. *Behavioral Genetics,* 5th edn. New York NY: Worth.

Plomin, R. & Kovas, Y. 2005. Generalist genes and learning disabilities. *Psychological Bulletin* 131: 592–617.

Plomin, R., Owen, M.J. & McGuffin, P. 1994. The genetic basis of complex human behaviors. *Science* 264: 1733–1739.

Rimol, L.M., Agartz, I., Djurovic, S., Brown, A.A., Roddey, J.C., Kähler, A.K. et al. 2010. Sex-dependent association of common variants of microcephaly genes with brain structure. *Proceedings of the National Academy of Sciences* 107: 384–388.

Scerif, G. & Karmiloff-Smith, A. 2005. The dawn of cognitive genetics? Crucial developmental caveats. *Trends in Cognitive Sciences* 9: 126–135.

Schrödinger, E. 2000/1944. *What is life*? Cambridge MA: Cambridge University Press: Canto.
Stromswold, K. 2001. The heritability of language: A review and meta-analysis of twin and adoption studies. *Language* 77: 647–723.
Toga, A.W. & Thompson, P.M. 2005. Genetics of brain structure and intelligence. *Annual Review of Neuroscience* 28: 1–23.
Tomasello, M. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge MA: Harvard University Press.
Tomblin, J.B. & Christiansen, M.H. 2009. Explaining developmental communication disorders. In *Speech Sound Disorders in Children: In Honor of Lawrence D. Shriberg*, R. Paul & P. Flipsen (eds), 35–49. San Diego CA: Plural Publishing.
Topol, E.J., Smith, J., Plow, E.F. & Wang, Q.K. 2006. Genetic susceptibility to myocardial infarction and coronary artery disease. *Human Molecular Genetics* 15(supplement 2): R117-123.
Tuomilehto-Wolf, E., Tuomilehto, J., Hitman, G.A., Nissinen, A., Stengård, J., Pekkanen, J. et al. 1993. Genetic susceptibility to non-insulin dependent diabetes mellitus and glucose intolerance are located in HLA region. *British Medical Journal* 307: 155–159.
Vernes, S.C., Newbury, D.F., Abrahams, B.S., Winchester, L., Nicod, J., Groszer, M. et al. 2008. A functional genetic link between distinct developmental language disorders. *The New England Journal of Medicine* 359: 2337–2345.
Walters, J.T.R. & Owen, M.J. 2007. Endophenotypes in psychiatric genetics. *Molecular Psychiatry* 12: 886–890.
Wheeler, M. & Clark, A. 1999. Genic representation: Reconciling content and causal complexity. *British Journal for the Philosophy of Science* 50: 103–135.
Woods, R.P., Freimer, N.B., De Young, J.A., Fears, S.C., Sicotte, N.L., Service, S.K. et al. 2006. Normal variants of Microcephalin and ASPM do not account for brain size variability. *Human Molecular Genetics* 15: 2025–2029.