NIH-PA Author Manuscript

Genetic Effects on Children's Conversational Language Use

Laura S. DeThorne
University of Illinois at Urbana–Champaign

Stephen A. Petrill and **Sara A. Hart**
The Ohio State University, Columbus

Ron W. Channell,
Brigham Young University, Provo, UT

Rebecca J. Campbell,
University of North Carolina, Chapel Hill

Kirby Deater-Deckard,
Virginia Polytechnic Institute and State University, Blacksburg

Lee Anne Thompson, and
Case Western Reserve University, Cleveland, OH

David J. Vandenbergh
The Pennsylvania State University, University Park

Abstract

Purpose—The present study examined the extent of genetic and environmental influences on individual differences in children's conversational language use.

Method—Behavioral genetic analyses focused on conversational measures and 2 standardized tests from 380 twins ($M = 7.13$ years) during the 2nd year of the Western Reserve Reading Project (S. A. Petrill, K. Deater-Deckard, L. A. Thompson, L. S. DeThorne, & C. Schatschneider, 2006). Multivariate analyses using latent factors were conducted to examine the extent of genetic overlap and specificity between conversational and formalized language.

Results—Multivariate analyses revealed a heritability of .70 for the conversational language factor and .45 for the formal language factor, with a significant genetic correlation of .37 between the two factors. Specific genetic effects were also significant for the conversational factor.

Conclusions—The current study indicated that over half of the variance in children's conversational language skills can be accounted for by genetic effects with no evidence of significant shared environmental influence. This finding casts an alternative lens on past studies that have attributed differences in children's spontaneous language use to differences in environmental language exposure. In addition, multivariate results generally support the context-dependent construction of language knowledge, as suggested by the theory of activity and situated cognition (J. S. Brown, A. Collins, & P. Duguid, 1989; T. A. Ukrainetz, 1998), but also indicate some degree of overlap between language use in conversational and formalized assessment contexts.

Contact author: Laura S. DeThorne, University of Illinois at Urbana–Champaign, 901 South Sixth Street, Champaign, IL 61820. E-mail: lauras@uiuc.edu.

A number of twin studies have used quantitative genetic methods to estimate environmental and genetic influences on language development. The twin design hinges on a comparison of monozygotic (MZ) and dizygotic (DZ) twins. Because MZ twins share 100% of their segregating genes and DZ twins share on average 50%, higher similarity between MZ twins is indicative of genetic effects (Plomin, DeFries, McClearn, & McGuffin, 2001). One means of measuring twin similarity is by comparing intraclass correlations for MZ versus DZ twins. The larger the MZ intraclass correlation in comparison with the DZs, the higher heritability (h^2) will be. In contrast, shared environmental effects (c^2) lead to similarity across all twins. Consequently, similar intraclass correlations between MZ and DZ twins are indicative of shared environmental effects. Finally, the extent to which MZ twins appear dissimilar is attributed to a combination of nonshared environment and error (e^2). Nonshared environmental influences are unique to the individual.

An underlying assumption of twin methodology is that the nature of genetic and environmental effects on the trait of interest—in this case, language ability—is not different between twins and singletons. Although twins appear at risk for slower language development (Conway, Lytton, & Pysh, 1980; Hay, Prior, Collett, & Williams, 1987; Rutter, Thorpe, Greenwood, Northstone, & Golding, 2003), such risk is associated with environmental influences, such as prematurity and low birth-weight, that are risk factors for singleton language delay as well (Field, Dempsey, & Shuman, 1981; Luke & Keith, 1992). In sum, although twins may be at increased risk for language difficulties, especially in early development, there is scarce indication that the cause of their language difficulties is qualitatively different than in singletons. Consequently, generalization appears warranted.

Based on a review of twin studies conducted by Plomin and Kovas (2005), estimates of genetic effects on child language abilities have varied widely, ranging from 16% to 100% (p. 595). The wide range of estimates is likely due to a number of factors, including variability in sample characteristics, such as child age, and in the domain of language being studied. For example, there is some evidence to suggest that genetic effects may increase as children get older, with shared environment accounting for less individual variation (e.g., Spinath, Price, Dale, & Plomin, 2004). In addition, a meta-analysis of twin studies by Stromswold (2001) suggested that genetic effects on language may vary by domain, with syntactic abilities associated with higher heritability than measures of lexical abilities.

In addition to child characteristics and language domain, the method of assessment may influence estimates of genetic and environmental effects. For example, the use of parent-report measures could inflate estimates of familiality by filtering assessments of both twins through the same observer (see Bishop, Laws, Adams, & Norbury, 2006). In addition to such methodological concerns, there are theoretical reasons to suspect that etiological influences on children's language skills may vary based on the method of assessment being used. The theory of activity and situated cognition (TASC; Brown, Collins, & Duguid, 1989; Ukrainetz, 1998) highlights the importance of context on mental representation and states that skills such as language are not learned independently from the context in which they were acquired. Consequently, one might predict that the causes of individual variance in the language skills used in conversation may not overlap with the causal factors of language used within the paradigm of a formalized test. Such contexts offer different cues (e.g., semantic, syntactic, visual) and inspire largely different motives (e.g., communication vs. performance). One might envision that conversational language skills are stored within a mental network that draws upon a child's degree of extraversion, pragmatic abilities, and skills at using contextual linguistic cues, such as semantic and syntactic bootstrapping (Hirsh-Pasek & Golinkoff, 1996). In contrast, formalized assessments represent a decontextualized schemata for language and may require children to draw on a distinct set of skills and cues, such as attention, motivation, frustration tolerance, and use of test-taking strategies (Dreisbach & Keogh, 1982; Erickson,

1972; Peña, Iglesias, & Lidz, 2001; Speltz, DeKlyen, Calderon, Greenberg, & Fisher, 1999). Given such context-related effects, one cannot assume that quantitative genetic findings from standardized tests or parent-report measures will generalize to children's spontaneous language use within a conversational context.

To date, studies of the genetic and environmental influences on language ability have focused on the use of structured probes, parent-report measures, or standardized tests and have not been extended to direct assessments of children's conversational language use on a large scale (e.g., Bishop, North, & Donlan, 1995; Dale et al., 1998; Eley et al., 1999; Eley, Bishop, Dale, Price, & Plomin, 2001; Spinath et al., 2004; Tomblin & Buckwalter, 1998). Instead, studies of children's conversational language have largely used correlation or group designs of parent–child interaction (e.g., Hart & Risley, 1995; Paul & Elwood, 1991), which do not control for potential genetic influences on language use. For example, associations between child language ability and amount of language exposure could result from innate differences in child language ability, which in turn elicit differences in environmental exposure. Extending behavioral genetic findings to children's language use in conversation is critical to teasing apart such relationships, particularly given the inherent social validity of this context. It is through conversation that individuals express needs, ideas, interests, and desires—thereby shaping their environments and their relationships with others.

We could find only two twin studies that directly assessed children's conversational language use, and neither study derived estimates of genetic and environmental effects. Stromswold and Rifkin (1996) collected longitudinal speech samples from two pairs of twins in which all 4 children were diagnosed with specific language impairment: one MZ pair and one DZ pair. The authors noted that in all measures, including mean length of utterance (MLU), the MZ twins were more similar to one another than the DZ twins throughout the course of the longitudinal study. Specifically, the MLU intraclass correlation was .92 for the MZ twins and .36 for the DZ twins. In contrast, Mather and Black (1984) calculated intraclass correlations for MLU and developmental sentence score (DSS; Lee, 1974) as well as standardized language measures from 158 preschool twins. The authors noted that only 71 children generated enough utterances in their language samples to obtain reliable measures. The resulting estimates of MZ and DZ correlations were, in part, counterintuitive: .60 and .90, respectively, for DSS and .44 and .46, respectively, for MLU. Although the similar intraclass correlations for MLU could be interpreted as support for shared environmental effects, it is hard to interpret the DZ correlation being higher than the MZ correlation for DSS. Greater DZ than MZ resemblance cannot be explained by the additive genetic model employed in most twin studies. The authors themselves offered no explanation and noted that the analyses of MLU and DSS are "included only for interest; they cannot provide evidence for the influence of heredity and environment on these skills" (Mather & Black, 1984, p. 305) due to the limited number of analyzable samples. An additional outcome of interest from Mather and Black (1984) related to the issue of assessment context. Specifically, a reported factor analysis revealed two dimensions across the children's language measures: one encompassing all the standardized tests and one encompassing their nonstandardized measures, including MLU and DSS. This finding is consistent with the TASC and highlights the need to derive genetic and environmental estimates for language measures across different contexts.

The present study was intended to address the paucity of behavioral genetic studies in the area of children's conversational language use, with a particular interest in estimates across measurement contexts. Specifically, our questions were as follows:

1.   To what extent are individual differences in children's conversational language use influenced by genetic versus environmental factors?

2. What is the extent of causal specificity and overlap across language measures from different assessment contexts (i.e., conversation vs. formalized assessment)?

Method

Participants

This study emerged from the Western Reserve Reading Project (WRRP; Petrill et al., 2006), a longitudinal twin study of gene-environment processes in the development of reading, mathematics, and related cognitive abilities. All twins were recruited primarily from Ohio, with the majority of families living in the metropolitan areas of Cleveland, Columbus, and Cincinnati. The present study began with 222 same-sex twin pairs with complete data in regard to zygosity, sex, age, hearing status, and transcribed conversational samples. Of those 222 pairs, 14 pairs were excluded because at least one of the twins within the pair had a history of hearing difficulty based on caregiver report. An additional 18 pairs were excluded because the conversational sample collected from one or both of the twins was considered too short (fewer than 50 complete and intelligible independent clauses) to provide reliable output measures (cf. Gavin & Giles, 1996). The remaining sample of 380 twins included 78 MZ pairs (38% male, 62% female) and 112 DZ pairs (44% male, 56% female). The mean age of the total sample was 7.13 years ($SD = .65$), with the mean age for MZs being 7.14 ($SD = .72$) and the mean age for DZs being 7.13 ($SD = .61$). The group mean on the Stanford–Binet Intelligence Scale (Thorndike, Hagen, & Sattler, 1986) was 101.81 ($SD = 12.69$), with the MZ and DZ means being identical and accompanied by SDs of 12.56 and 12.81, respectively. Based on caregiver questionnaire data, 12% of the twins were reported to have difficulties in expressive language (i.e., vocabulary and/or grammar), and 7% of the twins were being seen by a speech-language pathologist at the initiation of the WRRP (see DeThorne et al., 2006, for additional details).

Based on caregiver self-report of race/ethnicity, the sample was 93% White, 5% African American, 1% Asian, 1% Hispanic, and 1% other (percentages exceed 100% due to rounding error). All families reported using English in the home. In terms of the highest level of education attained by the child's primary caregiver, approximately 1% of the sample reported having no high school diploma, 7% reported having a high school diploma or equivalent, 16% reported having some college education, 10% completed a 2-year college degree, 32% completed 4 years of college, 6% reported having attended some graduate or professional school, 23% completed graduate or professional school, and 5% reported "other." Twin types were relatively similar in terms of parental education, with all the parents of MZ twins and all but 1% of the parents of DZs having received a high school diploma or equivalent.

General Procedures

Twins from the WRRP were recruited after they entered kindergarten but before they finished first grade and were assessed four times during the initial 3-year period of the project via home visit and parent/teacher questionnaires. Twins were recruited through media advertisements, school nominations, Ohio state birth records, and mothers of twins clubs. Once identified, participating families were mailed initial questionnaires to obtain information regarding demographics, pre/perinatal care, home environmental/parental attitudes, and history of speech-language development. After the initial questionnaires were returned, families were contacted to schedule an initial home visit, which focused on the assessment of early reading and factors related to the home environment. The second home visit, scheduled approximately 1 year after the initial home visit, contained a similar assessment to Year 1 but included the collection of a conversational language sample. During all home visits, twins were assessed simultaneously in separate rooms by two different examiners who were thoroughly trained in the assessment protocol and were experienced in interacting with school-aged children. Twin zygosity was determined through buccal swabs collected at the initial home visit and sent to

The Pennsylvania State University for DNA testing. For the handful of families who did not consent to DNA testing, twin zygosity was determined via a measure of twin physical similarity reported to be 95% accurate when compared with DNA analyses (Goldsmith, 1991; Price et al., 2000).

Language Sample Collection

The present study focused on conversational language data collected during the second-year home visit. Each sample consisted of a 15-min conversational exchange between child and examiner while the two were engaged in a play activity with modeling clay. All examiners were trained in language sample collection procedures using general guidelines provided by Leadholm and Miller (1992) such as "Limit requests, directions, and closed-questions"; "Listen to and follow the child's interests"; and "Use open-ended questions when appropriate." Potential conversational topics included, but were not limited to, (a) TV programs, (b) school activities, (c) family and siblings, (d) relevant holidays, (e) sports, and (f) food and meals. The entire conversational exchange was recorded onto audiocassette using a Marantz analogue tape recorder (PMD201) and a SONY stereo digital omnidirectional microphone (ECM-717).

Language Sample Transcription

The audiotaped recordings of the conversational samples were sent to the first author's laboratory for transcription of both examiner and child utterances by trained research assistants. Transcription training, overseen by the first author, included completing the Systematic Analysis of Language Transcripts (SALT; Version 8.0; Miller, 2004) tutorial, reviewing a lab transcription manual, and transcribing two or more practice transcripts. Research assistants did not begin transcribing independently until they demonstrated utterance and morpheme reliabilities of 85% or better.

Samples from within twin pairs were transcribed by different research assistants who were unaware of twin zygosity. During transcription, the samples were coded according to SALT conventions, including guidelines for utterance boundaries, word spellings, bound morphemes, and mazes. Because of the age of our sample, additional guidelines for determining utterance boundaries were taken from Nippold's (1998) description of Communication units (C-units; see also Loban, 1976). Specifically, in addition to using intonation pattern, degree of semantic contingency, and length of pause to determine the segmentation of utterances, any independent clauses joined by coordinating conjunctions (i.e., *and, or, but*) were segmented. For example, the following sentence would be segmented into two C-units at the second conjunction: "Like in the winter we would go outside and do ice skating on the pond/or we would go sled riding down these really big hills." After the original transcription, each language sample was reviewed by a second research assistant who corrected any obvious errors (e.g., miscoded bound morphemes, misspelled words). Transcription reliability on 38 randomly selected transcripts (i.e., 10% of the sample) revealed a mean agreement of 92% ($SD = 0.05$) for C-unit boundaries and 91% ($SD = 0.04$) for individual morphemes.

Language Sample Measures

Six variables reflecting the structure and content of spoken language were derived for each child sample: mean length of C-unit (MLU-C), number of total words (NTW), number of different words (NDW), measure D, total number of conjunctions (TNC), and DSS.

MLU-C—Included as a global measure of expressive language use, MLU-C was derived via SALT using all complete and intelligible child C-units within the sample. Validity of utterance length as a meaningful measure of language development is supported by (a) developmental change during school-age years (e.g., Leadholm & Miller, 1992; Rice, 2004; Rice, Redmond, & Hoffman, 2006), (b) differentiation of language ability (e.g., Klee, Schaffer, May,

Membrino, & Mougey, 1989), and (c) frequent use in language research on school-age children (e.g., Leonard & Finneran, 2003; Rice, Wexler, & Hershberger, 1998; Swanson, Fey, Mills, & Hood, 2005).

NTW—Usually referred to as a measure of volubility or general language proficiency (e.g., Leadholm & Miller, 1992), NTW was calculated via SALTon 100 complete and intelligible child C-units. Support regarding the validity of NTW as an index of expressive language proficiency includes developmental change during the school-age years (Leadholm & Miller, 1992) and high correlation with other language sample measures, including utterance length and NDW (Ukrainetz & Blomquist, 2002; DeThorne, Johnson, & Loeb, 2005).

NDW—NDW has been widely used as a measure of semantic diversity or productive vocabulary size (e.g., Leadholm & Miller, 1992; Bornstein, Haynes, Painter, & Genevro, 2000). In the present study, NDW included a count of all different root words and was derived via SALT on the first 100 complete and intelligible child utterances. Validity evidence for NDW comes from (a) documented developmental change during the school-age years (Leadholm & Miller, 1992; Miller, Freiberg, Rolland, & Reeves, 1992), (b) correlation with standardized vocabulary measures (Ukrainetz & Blomquist, 2002), and (c) differentiation of child language ability (DeThorne & Watkins, 2006; Watkins, Kelly, Harbers, & Hollis, 1995; Miller, 2001).

Measure D—Measure D was included as a second measure of expressive vocabulary. Measure D has the advantage of minimizing the influence of sample size while using the entirety of the language transcripts, regardless of varying lengths (Malvern & Richards, 1997; McKee, Malvern, & Richards, 2000; Owen & Leonard, 2002). Specifically, measure D is derived for any particular transcript by calculating a series of type token ratios (TTRs) from hundreds of randomly selected subsamples of the transcript, each ranging from 35 to 50 tokens. The resulting curve is then compared with a theoretical distribution of curves that is based on a model of the relation between TTR and number of tokens. The theoretical curve that best fits the actual curve derived from the real transcript provides the specific measure D. We derived measure D through a program known as Vocd, available through the Child Language Data Exchange System (http://childes.psy.cmu.edu). Although relatively new, measure D has demonstrated change with age (Durán, Malvern, Richards, & Chipere, 2004; Klee, Stokes, Wong, Fletcher, & Gavin, 2004; Owen & Leonard, 2002), correlated with other measures of expressive vocabulary (Malvern & Richards, 2002; Wright, Silverman, & Newhoff, 2003), and successfully differentiated groups (e.g., Durán et al., 2004; Owen & Leonard, 2002; Wright, Silverman, & Newhoff, 2003).

TNC—TNC was included as a measure of complex syntax (Nippold, 1998) that has demonstrated a developmental trend during the school-age years (Leadholm & Miller, 1992). TNC was calculated on 100 complete and intelligible child C-units and included a frequency count of both subordinating and coordinating conjunctions, specifically the following 12 types: *after, and, as, because, but, if, or, since, so, then, until,* and *while*.

DSS—As an additional measure of morphosyntactic complexity, DSS was calculated using all eligible child utterances in a sample. In addition to being complete and intelligible, eligible utterances had to contain both a subject and a predicate (Lee, 1974). Because of concerns regarding reliability, DSS was not calculated for transcripts with fewer than 35 eligible child utterances (Johnson & Tomblin, 1975).

DSS reflects children's morphosyntactic development in eight different categories: (a) indefinite pronouns/noun modifier, (b) personal pronouns, (c) main verb forms, (d) secondary verb forms, (e) negative constructions, (f) conjunctions, (g) interrogative reversals in

questions, and (h) *wh*-question forms. DSS has revealed group differences between typically developing children and those with language impairment (e.g., Johnston & Kamhi, 1984; Rice, Redmond, & Hoffman, 2006) and has served to document treatment effects on language development (e.g., Hughes, Fey, & Long, 1992).

After utterance selection, tentative DSS codes were automatically derived using a software program titled gcSalt (Channell, 2006), a program that parallels DSS in terms of scope and function. Afterward, the automated scores were manually corrected by the fourth author according to DSS coding procedures specified in Lee (1974) and clarified by Lively (1984), with two procedural exceptions. First, the scoring of "like" was changed, as suggested by Hughes, Fey, and Long (1992). Second, as utterances had been transcribed as subsentential C-units, all initial conjunctions (except "and") were scored.

Formalized Vocabulary Measures

Two formalized assessments of vocabulary were included in the present study to examine the issue of causal overlap and specificity with the conversational measures. Specifically, the twins completed the Boston Naming Test (BNT; Goodglass & Kaplan, 2001), which required them to name a series of pictured objects, and the Vocabulary subtest from the Stanford–Binet Intelligence Scale (SB-Vocab; Thorndike, Hagen, & Sattler, 1986), which required them to define words.

Analyses

Research Question 1, regarding the extent of genetic and environmental influences on children's conversational language use, was addressed through intraclass correlations (Falconer, 1960) and structural equation modeling, which provides inferential testing of heritability (h^2), shared environmental (c^2), and nonshared environmental (e^2) estimates (Neale, Boker, Xie, & Maes, 1999). Each estimate represents a portion of the total individual variance, and together all estimates add up to 1. Estimates are statistically significant when the related confidence interval does not encompass zero.

Research Question 2, regarding the extent of specificity and overlap across measures from different contexts, was addressed through correlation, latent trait analyses, and multivariate genetic modeling. Correlation analyses addressed phenotypic associations, whereas multivariate modeling examined the genetic and environmental contribution to specificity and overlap of latent factors (e.g., conversational vs. formal). Described in more detail in the Results section, the multivariate genetic analyses conceptually mirror the univariate analyses. Instead of comparing scores from twins and co-twins on the same measure, twins' performance on one factor is compared with their co-twins' performance on a second factor. As in the univariate analyses, if the covariance is higher in MZ twins than in DZ twins, genetic effects are implicated.

Results

Descriptives

Descriptive data on the language sample measures are summarized in Table 1, with values provided separately for MZ and DZ twins. The conversational language samples ranged from 50 to 272 complete and intelligible C-units, with a mean of 144.32 (*SD* = 42.30). Recall that twin pairs were originally excluded from this study if either twin had fewer than 50 complete and intelligible C-units. Data on DSS were restricted to the 320 children whose samples contained a minimum of 35 eligible utterances for both twins. Visual inspection of the entire twin sample revealed relatively normal distributions for all language sample measures except TNC, with skewness values between −.350 and .297 and kurtosis values ranging from .208 to .

687. The distribution for TNC was positively skewed (skewness = .996) and leptokurtic (kurtosis = 1.36). To test assumptions in the twin design, Levene's test for equality of variances was employed. Variance differences emerged for four of the six measures, with the DZ group demonstrating greater variability than the MZ group in each case: MLU-C ($F = 7.81$, $p = .005$), NTW ($F = 7.98$, $p = .005$), NDW ($F = 4.69$, $p = .031$), and TNC ($F = 5.51$, $p = .019$). The potential impact of such differences is addressed in the Discussion section. Using an alpha of .05, no significant mean differences emerged between MZ twins and DZ twins for any of the language sample measures.

Correlation Analyses

All the child language sample measures except measure D correlated significantly with age at an alpha of .001. Coefficients ranged from .194 for NTW to .294 for DSS, with the percentage of variance accounted for being small (Cohen, 1988). To better understand the nature of the measures used, a partial correlation matrix was derived for all language measures, controlling for child age (see Table 2). Alpha was set at .002 (.05/28) using a Bonferroni adjustment to help control for spurious significance. All conversational measures correlated significantly except measure D with MLU-C, NTW, and DSS.

Group Comparisons

To address the validity of our language sample measures, *t* tests were conducted to compare mean values for children reported by their caregiver to have expressive language difficulties at the initiation of the study ($n = 45$) versus those reported not to have such difficulties ($n = 333$). The incidence of reported difficulties was 18% in MZ twins (27 of 154) and 8% in DZ twins (18 of 224). Language sample values were age- and sex-corrected using regression procedures before being evaluated for group differences. Children with a history of expressive language difficulties scored significantly lower than children without a history of expressive difficulties on all six measures: MLU-C ($t = 4.16$, $df = 376$, $p = .000$), NTW ($t = 3.76$, $df = 365$, $p = .000$), NDW ($t = 4.73$, $df = 365$, $p = .000$), TNC ($t = 3.10$, $df = 365$, $p = .002$), measure D ($t = 2.25$, $df = 376$, $p = .025$), DSS ($t = 4.27$, $df = 316$, $p = .000$). Mean differences ranged from approximately 1/3 of a standard deviation in the case of measure D to nearly 3/4 of a standard deviation in the case of DSS and NDW.

Univariate Genetic Analyses

Using the age- and sex-corrected values, intraclass correlations for MZ and DZ twins were derived for initial information regarding evidence of genetic and/or shared environmental influences. As can be seen in Table 3, the significant MZ intraclass correlations provided evidence of familiality across all measures, with DZ correlations being significant as well for MLU-C, NTW, NDW, TNC, and the BNT. To estimate and test the significance of genetic and environmental effects, univariate ACE structural equation models were applied to raw data comprised of the age- and sex-corrected language values (Neale & Cardon, 1992). Results are presented in Table 3. The heritability estimates for MLU-C, NTW, NDW, and the BNT reached statistical significance. Estimates of shared environmental effects were significant only for the BNT.

Note that the univariate estimates appear somewhat inconsistent with the intraclass correlations. For example, an informal rule of thumb for estimating heritability is to double the difference between MZ and DZ intraclass correlations. In the case of MLU-C, this procedure would produce an estimated heritability of .27 compared with the modeling estimate of .49. Such inconsistency is due to the variance differences observed across MZ and DZ twins. Power limitations also need to be noted here. Based on Mx power analyses, we would have needed approximately 500 to 1,000 more twin pairs to reach statistical significance with the observed effects sizes for measure D, TNC, DSS, and the Vocabulary subtest of the Stanford–Binet

Intelligence Test (SB-Vocab). The issues of variance differences and power limitations will be addressed further in the Discussion section.

Multivariate Genetic Analyses

To address the issue of specificity and overlap at an etiologic level, we created two factors based on latent traits and then decomposed the covariance and variance of the two factors into genetic and environmental influences (see Figure 1). Given our interest in the contrast between conversational and formal testing contexts, we assigned the two formalized assessments, SB-Vocab and BNT, to Factor 1 (hereby referred to as the *formal factor*), and we assigned the conversational measures, MLU, NTW, NDW, measure D, TNC, and DSS, to Factor 2 (hereby referred to as the *conversational factor*). Factor loadings were acceptable, with variables for the formal factor associated with loadings of .52 and .93 for SB-Vocab and BNT, respectively. The residuals were .73 and .14 accordingly, representing the variance not accounted for by the formal factor. The conversational factor was associated with factor loadings ranging from .65 to .97 with the exception of measure D, which demonstrated the weak loading of .19. The residuals ranged from .05 to .96.

In addition to the measurement solution, the model simultaneously used a bivariate Cholesky decomposition to examine the genetic and environmental contributions to the variance and covariance between the two factors. As presented in Figure 1, the first set of estimates, genetic (A_1), shared environment (C_1), and unique environment (E_1), measure the overlap between the formal and conversational factors. The second set of estimates (i.e., A_2, C_2, and E_2) measures the unique genetic and environmental influences of the conversational factor alone (Neale & Cardon, 1992). Results, as seen in Table 4, revealed significant genetic etiology for both formal ($h^2 = .67^2$, or .45) and conversational ($h^2 = .31^2 + .78^2$, or .70) factors. Shared environmental effects were also significant for the formal factor ($c^2 = .74^2$, or .55) but not for the conversational factor. Child-specific nonshared environmental influences approached zero for the formal factor ($e^2 = .04^2$, or .001) and were nonsignificant for the conversational factor ($e^2 = .55^2$, or .30), as well.

This model also tested the genetic and environmental etiology of the overlap between the formal and conversational factors. In particular, statistically significant genetic overlap is implied by significant path estimates for the A_1 factor for both formal and conversational factors. The second genetic factor (A_2), which represents the unique genetic influences on the conversational factor independent from those attributable to the overlap with the formal factor, was also significant with a path estimate of .78. Consistent with the univariate results, there was no evidence of shared or unique environmental overlap between the two factors or on the conversational factor alone.

Put another way, the estimated phenotypic correlation between formal and conversational factors is cal culated through the sum of the products of the A_1, C_1, and E_1 factor loadings (i.e., $.67 \times .31 + .74 \times .00 + .04 \times .55 = .23$). Given the phenotypic correlation of .23, .21 ($.67 \times .31$) is attributable to genetic covariance between the formal and conversational factors. Thus, although the phenotypic correlation between the formal factor and conversational factor is modest, 91% (.21/.23) of that correlation is due to genetic covariance. In addition to the overlap, there also appears to be considerable genetic independence for conversational language. Specifically, the significant A_2 factor of .78 indicates that 60% of the variance in conversational language is not only genetically influenced but also independent from the formal language factor.

Discussion

The present study indicated genetic effects on children's formal and conversational language use. In addition, there was significant genetic overlap between latent factors of formal and conversational language as well as significant specific genetic effects for the conversational factor. Before discussing the implications of these findings, a brief review of issues related to internal and external validity is addressed.

In regard to internal validity, we will highlight four forms of evidence supporting the conversational measures as meaningful indices of children's expressive language use, with the possible exception of measure D. First, all language sample measures, except measure D, correlated positively with child age, and all approximated normal distributions except for TNC. The positively skewed distribution of TNC likely reflected the use of C-units for utterance segmentation. Specifically, the nature of C-units limits the number of conjoining conjunctions that can be included in an utterance, leading to a positively skewed distribution. The second form of validity evidence is that the conversational language measures were similar to reference data published for typically developing children in a comparable age group. For example, 64 children from the WisconsinConCunits database in SALT (Version 8), ranging in age from 6;1 [years;months] to 8;9, generated a mean MLU of 5.26 (*SD* = 0.74), which approximates the sample means of 5.72 and 5.68 for MZ and DZ twins, respectively, within this study. Group means for NTW, NDW, and TNC were all within 1 *SD* of the SALT reference values, as well. Of interest, this finding calls into question the common notion that twins tend to have less advanced language skills compared with singletons. Perhaps reported delays are most pronounced at younger ages (Conway et al., 1980; Hay et al., 1987; Rutter et al., 2003). The third form of evidence supporting the validity of the language sample measures is the consistent mean differences favoring children without a history of expressive language difficulties compared with those who have a reported history of expressive difficulties. Last, all measures except measure D loaded strongly on a single factor that we have called *conversational language*.

Together, these findings suggest that our language sample measures are largely consistent with prior studies and are sensitive to meaningful differences in the language skills of our sample. The possible exception is that measure D, which failed to correlate with child age, demonstrated the smallest mean difference across children grouped by language ability and failed to load strongly on the conversational language factor. Although the reason for such findings is unclear, we have hypothesized that expressive vocabulary ability and talkativeness are inextricably linked within the conversational context and by attempting to control for the latter, the calculation of measure D may actually remove a large portion of meaningful variance in children's conversational language use (DeThorne, Coletto, Wendorf, Petrill, & Johnson, 2007).

In addition to accurate assessment, internal validity is contingent on meeting methodological assumptions. Twin methodology is built on the assumption of equal means and variances across MZ and DZ groups. Although group means were equal across all language sample measures, variance differences were observed for MLU-C, NTW, NDW, and TNC, with less variance observed in MZ twins for all variables. The source of this inequality is unknown. However, when the larger group reflects the larger variance, which is the case here, the violation of equal variances leads to more conservative estimates of effects and helps explain why results from the intraclass correlations and model estimates appear somewhat discrepant. Also of concern is the statistical power associated with our univariate analyses, which makes it possible that observed trends in the univariate analyses toward heritability for measure D, TNC, DSS, and SB-Vocab may represent meaningful effects that failed to reach significance and are worth future study. In sum, unequal variances and limited power in the present study may have

increased our likelihood of Type II errors but does not invalidate the observed significant findings.

Given support for the internal validity of our study, we turn now to a discussion of external validity, or the extent to which results can be expected to generalize to other populations. We already discussed support for the generalization from twin research to the singleton population, so we focus here on other important sample characteristics. For example, although our participating families were recruited to be representative of the larger U.S. population, it appears to underrepresent parents with less education. Because parent education may correlate with relevant causal influences, be they genetic or environmental, it is possible that we underestimated the influence of such factors, and our results may not generalize to samples containing a larger percentage of parents who did not receive high school diplomas. Similarly, our inclusion of children with language disabilities may lead to different estimates of heritability than studies of typical language variation that exclude children with disability. There has been some evidence to suggest that the extent of heritability increases as children's abilities fall further below the population mean (e.g., DeThorne et al., 2006). Consequently, samples that exclude children with language disabilities might expect somewhat lower estimates of heritability. Finally, it is important to note that the present study did not include formalized measures of grammar. Consequently the observed genetic overlap between conversational and formalized language abilities may be limited to expressive vocabulary and may fail to generalize to other language domains.

Having addressed issues regarding internal and external validity, we now turn toward a comparison of present results to prior studies and theoretical predictions. In regard to prior literature, this study is the first to document the extent of genetic effects on direct assessments of children's conversational language. The magnitude of heritability is comparable to that reported for child language ability assessed through other methods (Plomin & Kovas, 2005; Stromswold, 2001). Similarly, the finding was consistent with trends reported in the longitudinal study of spontaneous language in two twin pairs reported by Stromswold and Rifkin (1996). In potential contrast, the intraclass correlations reported by Mather and Black (1984) were more consistent with shared environmental effects, although the authors themselves hesitate to draw conclusions from these correlations, given the small number of participants included. It is also worth noting that the children from Mather and Black's study were at a younger age, when estimates of shared environmental effects may be stronger.

Given our finding of genetic effects on children's conversational language, it becomes of interest whether or not the same genes are influencing children's language use across conversational and formalized assessment contexts. It is possible for two traits to be heritable to a similar extent and yet be influenced by entirely separate genes. The theory of activity and situated cognition (TASC; Brown, Collins, & Duguid, 1989; Ukrainetz, 1998) emphasizes the role of context in learning and highlights the possibility that the etiology of individual differences in conversational language may not overlap with the etiology of individual differences in more formalized language use. Of interest, results from the present study were somewhat mixed in this regard.

In line with the TASC, the latent trait analyses was consistent with two factors, one encompassing the conversational measures (except measure D) and one encompassing the formalized vocabulary measures. In addition, the conversational factor demonstrated a significant amount of independent genetic effects. However, two findings made interpretation in regard to the TASC more complex. First, measure D, a conversational measure of lexical diversity, did not load on the conversational factor as predicted, although it did correlate significantly with NDW. As previously mentioned, it is possible that the unique way in which D is calculated sets it apart from the other conversational measures. A second finding that

contradicts a strict interpretation of the TASC is the significant genetic overlap that emerged between the formalized and conversational factors. Specifically, the genetic correlation of .37 indicated that a significant amount of genetic variance was shared between the two factors. However, it is important to keep in mind that the phenotypic correlations across the formalized and conversational factors were relatively small in effect size (Cohen, 1988), with 60% of the variance in children's conversational language use due to genetic effects that were independent from the formal language factor.

In sum, results are generally supportive of the context-dependent construction of language knowledge as suggested by the TASC but also suggest the presence of some causal overlap with language used in conversational and formalized assessment situations. Our view of the TASC is not that it precludes any degree of causal overlap across contexts but that it predicts a certain degree of specific causal effects that are dependent on the context of the assessment. In this regard, the TASC prediction was realized. Overall, the extent of phenotypic and genetic overlap across language measures is likely impacted both by similarities/differences in the assessment contexts of the tasks (e.g., conversation vs. formal testing) as well as the domain of language being assessed (e.g., vocabulary vs. grammar), with the highest overlap predicted between measures of the same language domain sampled from the same context.

Implications

The current study indicated that differences in children's conversational language use can be accounted for, in part, by genetic effects. This finding casts an alternative lens on past studies that have correlated children's spontaneous language use with differences in environmental exposure (e.g., Hart & Risley, 1995). Whereas such correlations have been interpreted as differences in exposure causing individual differences in children's vocabulary growth, the present study highlighted the possibility that differential exposure may, in part, reflect a gene–environment correlation. For example, children with a genetic predisposition toward strong word learning skills may elicit more words from their caregivers (e.g., evocative gene–environment correlation) or may tend toward having parents with stronger vocabulary skills themselves who tend to use more words with their children (e.g., passive gene–environment correlation). This change in perspective may not change the general suggestion to facilitate children's exposure to language, but it does suggest that the probability of children being exposed to a "language-rich" environment may be, in part, genetically mediated. Targeted interventions may, in fact, provide an opportunity to break the link between genetic risk and at least one form of environmental risk. It is critical to highlight the fact that "genetic" is not synonymous with "unchangeable." In fact, understanding the complex pathway from genes to behavior will help us devise the most effective interventions and preventions, thereby allowing us to reduce or eliminate the translation of genetic risk into functional disability.

Finally, from a diagnostic and prognostic standpoint, information regarding family history of language difficulties is advantageous. For example, one could imagine cases in which a child's diagnostic information is "borderline" or the results from different assessments are conflicting. Information regarding the prevalence and severity of language difficulties in other members of a child's family, in conjunction with other factors, could help guide difficult clinical decisions regarding the initiation, continuation, or intensity of intervention.

Also in regard to assessment, one might wonder whether or not the genetic overlap between formalized and conversational measures implies that both types of measures are unnecessarily redundant within an assessment protocol. However, it is important to keep in mind that although the genetic overlap is significant, the phenotypic correlation is low, with over half of the variance in conversational language being explained by genetic factors that are independent from formalized language use. Consequently, genetic overlap is encouraging in regard to the

construct validity of these measures but does not suggest that assessment can rely on measurement from one context. Recommendations for best practice continue to focus on incorporating multiple forms of assessment, including standardized assessment, observational data such as conversational samples, and interviews (see Watkins & DeThorne, 2000).

In regard to future research, evidence of genetic effects on conversational language naturally leads to an interest in identifying relevant loci. This daunting task will likely require the collective implementation of multiple methodologies, including additional behavioral genetic work and molecular studies. The present results would predict that some, but not all, of identified loci associated with conversational language abilities would also be associated with psychometric expressive vocabulary abilities.

In addition to genetic overlap, evidence of specific nongenetic effects offers another avenue of interest. Intervention would profit from the identification of these influences and how they differentially impact one form of measurement versus the other. For example, the conversational partner may have a greater impact on children's conversational measures than the examiner has on children's formalized test scores because the examiner's role is more prescribed in the latter context. Consequently, identifying the nongenetic influences that improve a child's conversational language use but do not generalize to formalized test performance, and vice versa, is of interest. In sum, maximally effective efforts to facilitate children's language development hinge on the ability to identify not only the relevant genetic and environmental factors but also how each factor impacts children's language use in different contexts.

References

Bishop DVM, Laws G, Adams C, Norbury CF. High heritability of speech and language impairments in 6-year-old twins demonstrated using parent and teacher report. Behavior Genetics 2006;36:173–184. [PubMed: 16485179]

Bishop DVM, North T, Donlan C. Genetic basis of specific language impairment: Evidence from a twin study. Developmental Medicine and Child Neurology 1995;37:56–71. [PubMed: 7828787]

Bornstein MH, Haynes OM, Painter KM, Genevro JL. Child language with mother and with stranger at home and in the laboratory: A methodological study. Journal of Child Language 2000;27:407–420. [PubMed: 10967894]

Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. Educational Researcher 1989;17:32–41.

Channell, RW. gcSalt (Version 1.02) [Computer software]. Provo, UT: Brigham Young University; 2006.

Cohen, J. Statistical power analysis for the behavioral sciences. 2. Hillsdale, NJ: Erlbaum; 1988.

Conway D, Lytton H, Pysh F. Twin-singleton language differences. Canadian Journal of Behavioral Science 1980;12:264–271.

Dale PS, Simonoff E, Bishop DVM, Eley TC, Oliver B, Price TS, et al. Genetic influence on language delay in two-year-old children. Nature Neuroscience 1998;1:324–328.

DeThorne, LS.; Coletto, M.; Wendorf, L.; Petrill, SA.; Johnson, BW. Measuring conversational vocabulary: Can we separate the baby from the bath water?; 2007, June; Poster session presented at the Symposium on Research in Child Language Disorders; Madison, WI.

DeThorne LS, Hart SA, Petrill SA, Deater-Deckard K, Thompson LA, Schatschneider C, Davison MD. Children's history of speech-language difficulties: Genetic influences and associations with reading-related measures. Journal of Speech, Language, and Hearing Research 2006;49:1280–1293.

DeThorne LS, Johnson BW, Loeb JW. A closer look at MLU: What does it really measure? Clinical Linguistics & Phonetics 2005;19:635–648. [PubMed: 16147407]

DeThorne LS, Watkins RV. Language abilities and nonverbal IQ in children with language impairment: Inconsistency across measures. Clinical Linguistics & Phonetics 2006;20:641–658. [PubMed: 17342874]

Dreisbach M, Keogh BK. Testwiseness as a factor in readiness test performance of young Mexican-American children. Journal of Educational Psychology 1982;74:224–229.

Durán P, Malvern D, Richards B, Chipere N. Developmental trends in lexical diversity. Applied Linguistics 2004;25:220–242.

Eley TC, Bishop DVM, Dale PS, Oliver B, Petrill SA, Price TS, et al. Genetic and environmental origins of verbal and performance components of cognitive delay in 2-year-olds. Developmental Psychology 1999;35:1122–1131. [PubMed: 10442880]

Eley TC, Bishop DVM, Dale PS, Price TS, Plomin R. Longitudinal analysis of the genetic and environmental influences on components of cognitive delay in preschoolers. Journal of Educational Psychology 2001;93:698–707.

Erickson ME. Test sophistication: An important consideration. Journal of Reading 1972;16:140–144.

Falconer, DS. Introduction to quantitative genetics. Edinburgh, Scotland: Oliver & Boyd; 1960.

Field, T.; Dempsey, J.; Shuman, H. Developmental follow-up of pre-term and post-term infants. In: Sigman, M.; Friedman, S., editors. Preterm birth and psychological development. New York: Academic Press; 1981. p. 299-312.

Gavin WJ, Giles L. Sample size effects on temporal reliability of language sample measures of pre-school children. Journal of Speech and Hearing Research 1996;39:1258–1262. [PubMed: 8959610]

Goldsmith HH. A zygosity questionnaire for young twins: A research note. Behavior Genetics 1991;21:257–269. [PubMed: 1863259]

Goodglass, H.; Kaplan, E. The Boston Naming Test. 2. New York: Lippincott Williams & Wilkins; 2001.

Hart, B.; Risley, TR. Meaningful differences in the everyday experience of young American children. Baltimore: Paul H. Brookes; 1995.

Hay DA, Prior M, Collett S, Williams M. Speech and language development in preschool twins. Acta Geneticae Medicae et Gemeollologiae: Twin Research 1987;36:213–223.

Hirsh-Pasek, K.; Golinkoff, RM. The origins of grammar: Evidence from early language comprehension. Cambridge, MA: The MIT Press; 1996.

Hughes DL, Fey ME, Long SH. Developmental sentence scoring: Still useful after all these years. Topics in Language Disorders 1992;12:1–12.

Johnson MR, Tomblin JB. The reliability of developmental sentence scoring as a function of sample size. Journal of Speech and Hearing Research 1975;18:372–380.

Johnston J, Kamhi AG. The same can be less: Syntactic and semantic aspects of the utterances of language-impaired children. Merrill-Palmer Quarterly 1984;30:65–86.

Klee T, Schaffer M, May S, Membrino I, Mougey K. A comparison of the age–MLU relation in normal and specifically language-impaired preschool children. Journal of Speech and Hearing Disorders 1989;54:226–233.

Klee T, Stokes SF, Wong AMY, Fletcher P, Gavin WJ. Utterance length and lexical diversity in Cantonese-speaking children with and without specific language impairment. Journal of Speech, Language, and Hearing Research 2004;47:1396–1410.

Leadholm, BJ.; Miller, JF. Language sample analysis: The Wisconsin guide. Madison, WI: Wisconsin Department of Public Instruction; 1992. Analysis; p. 36-51.

Lee, L. Developmental sentence analysis. Evanston, IL: Northwestern University Press; 1974.

Leonard L, Finneran D. Grammatical morpheme effects on MLU: "The same can be less" revisited. Journal of Speech, Language, and Hearing Research 2003;46:878–888.

Lively M. Developmental sentence scoring: Common scoring errors. Language, Speech, and Hearing Services in Schools 1984;15:154–168.

Loban, W. Research Report No. 18. Urbana, IL: National Council of Teachers of English; 1976. Language development: Kindergarten through grade twelve.

Luke B, Keith LG. The contribution of singletons, twins, and triplets to low birth weight, infant mortality, and handicap in the United States. Journal of Reproductive Medicine 1992;37:661–666. [PubMed: 1432978]

Malvern D, Richards B. Investigating accommodation in language proficiency interview using a new measure of lexical diversity. Language Testing 2002;19:85–104.

Malvern, DD.; Richards, BJ. A new measure of lexical diversity. In: Ryan, A.; Wray, A., editors. Evolving models of language: Papers from the Annual Meeting of the British Association of Applied Linguists. Philadelphia: Multilingual Matters; 1997. p. 58-71.

Mather PL, Black KN. Hereditary and environmental influences on preschool twins. Developmental Psychology 1984;20:303–308.

McKee G, Malvern D, Richards B. Measuring vocabulary diversity using dedicated software. Literary and Linguistic Computing 2000;15:323–337.

Miller, J. Language sample analyses for monitoring intervention: A clinician's toolbox; 2001, November; Presentation at the American Speech-Language-Hearing Association Convention; New Orleans, LA.

Miller, J. Systematic Analysis of Language Transcripts (Research Version 8.0) [Computer software]. Madison, WI: University of Wisconsin; 2004.

Miller JF, Freiberg C, Rolland M, Reeves MA. Implementing computerized language sample analysis in the public schools. Topics in Language Disorders 1992;12:69–82.

Neale, M.; Boker, SM.; Xie, G.; Maes, HH. Mx statistical modeling. 5. Richmond, VA: Virginia Commonwealth University; 1999.

Neale, MC.; Cardon, LR. Methodology for genetic studies of twins and families. Dordrecht, the Netherlands: Kluwer Academic; 1992.

Nippold, MA. Later language development: The school-age and adolescent years. 2. Austin, TX: Pro-Ed; 1998.

Owen AJ, Leonard LB. Lexical diversity in the spontaneous speech of children with specific language impairment: Application of D. Journal of Speech, Language, and Hearing Research 2002;45:927–937.

Paul R, Elwood TJ. Maternal linguistic input to toddlers with slow expressive language development. Journal of Speech and Hearing Research 1991;34:982–988. [PubMed: 1749252]

Peña E, Iglesias A, Lidz CS. Reducing test bias through dynamic assessment of children's word learning ability. American Journal of Speech-Language Pathology 2001;10:138–154.

Petrill SA, Deater-Deckard K, Thompson LA, DeThorne LS, Schatschneider C. Reading skills in early readers: Genetic and shared environmental effects. Journal of Learning Disabilities 2006;39:48–55. [PubMed: 16512082]

Plomin, R.; DeFries, JC.; McClearn, GE.; McGuffin, P. Behavioral genetics. 4. New York: Worth; 2001.

Plomin R, Kovas Y. Generalist genes and learning disabilities. Psychological Bulletin 2005;131:592–617. [PubMed: 16060804]

Price TS, Eley TC, Dale PS, Stevenson J, Saudino K, Plomin R. Genetic and environmental covariation between verbal and nonverbal cognitive development in infancy. Child Development 2000;71:948–959. [PubMed: 11016558]

Rice, M. Growth models of developmental language disorders. In: Rice, ML.; Warren, SF., editors. Developmental language disorders: From phenotypes to etiologies. Mahwah, NJ: Erlbaum; 2004. p. 207-240.

Rice ML, Redmond SM, Hoffman L. Mean length of utterance in children with specific language impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. Journal of Speech, Language, and Hearing Research 2006;49:793–808.

Rice ML, Wexler K, Hershberger S. Tense over time: The longitudinal course of tense acquisition in children with specific language impairment. Journal of Speech, Language, and Hearing Research 1998;41:1412–1431.

Rutter M, Thorpe K, Greenwood R, Northstone K, Golding J. Journal of Child Psychology and Psychiatry 2003;44:326–341. [PubMed: 12635964]

Speltz ML, DeKlyen M, Calderon R, Greenberg MT, Fisher PA. Neuropsychological characteristics and test behaviors of boys with early onset conduct problems. Journal of Abnormal Psychology 1999;108:315–325. [PubMed: 10369042]

Spinath FM, Price TS, Dale PS, Plomin R. The genetic and environmental origins of language disability and ability. Child Development 2004;75:445–454. [PubMed: 15056198]

Stromswold K. The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. Language 2001;77:647–723.

Stromswold, K.; Rifkin, JI. Language acquisition by identical vs. fraternal SLI twins; 1996, June; Poster session presented at the Symposium for Research in Child Language Disorders; Madison, WI.

Swanson LA, Fey ME, Mills CE, Hood LS. Use of narrative-based language intervention with children who have specific language impairment. American Journal of Speech-Language Pathology 2005;14:131–143. [PubMed: 15989388]

Thorndike, RL.; Hagen, EP.; Sattler, JM. The Stanford–Binet Intelligence Scale IV: Guide for administering and scoring. Itasca, IL: Riverside Publishing; 1986.

Tomblin JB, Buckwalter PR. Heritability of poor language achievement among twins. Journal of Speech, Language, and Hearing Research 1998;41:188–199.

Ukrainetz TA. Beyond Vygotsky: What Soviet activity theory offers naturalistic language intervention. Journal of Speech-Language Pathology and Audiology 1998;22:164–175.

Ukrainetz TA, Blomquist C. The criterion validity of four vocabulary tests compared to a language sample. Child Language Teaching and Therapy 2002;18:59–78.

Watkins RV, DeThorne LS. Assessing children's vocabulary: From word knowledge to word learning potential. Seminars in Speech and Language 2000;21:235–245. [PubMed: 10958432]

Watkins RV, Kelly DJ, Harbers HM, Hollis W. Measuring children's lexical diversity: Differentiating typical and impaired language learners. Journal of Speech and Hearing Research 1995;38:1349–1355. [PubMed: 8747826]

Wright HH, Silverman SW, Newhoff M. Measures of lexical diversity in aphasia. Aphasiology 2003;17:443–452.
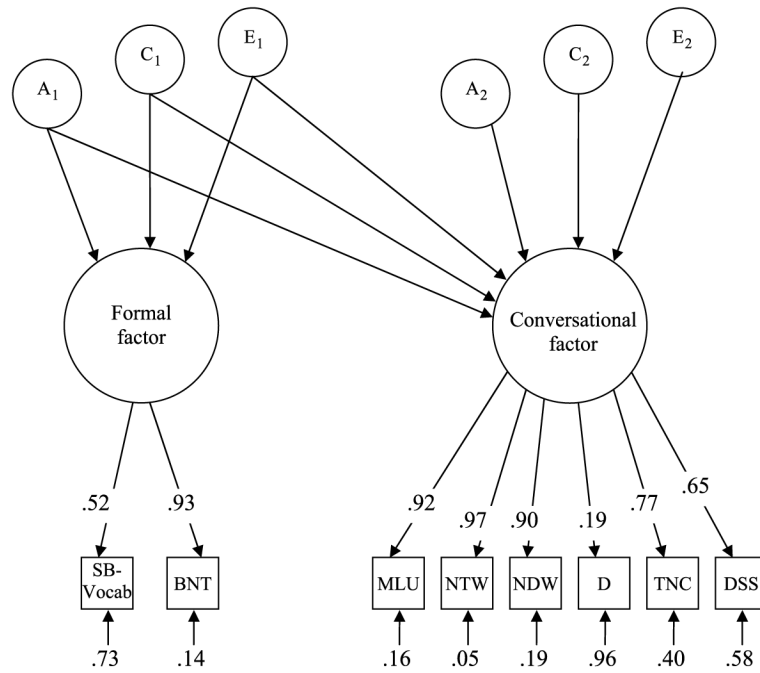
Figure 1.
Multivariate analyses using two latent factors. Note that the upper pathway estimates represent factor loadings for each individual measure and the lower values represent residuals (i.e., variance unexplained by the factor). The formalized factor includes the Vocabulary subtest from the Stanford–Binet Intelligence Test (Thorndike, Hagen, & Sattler, 1986) and the Boston Naming Test (Goodglass & Kaplan, 2001). The conversational factor includes mean length of C-unit, number of total words, number of different words, measure D, total number of conjunctions, and developmental sentence scores (Lee, 1974).

Table 1

Descriptive data on child language sample measures.

Child language variable	MZ			DZ		
	n	M (SD)	Range	n	M (SD)	Range
MLU-C	156	5.72 (.99)	2.88–8.43	224	5.68 (1.22)	2.65–9.50
NTW	151	504.41 (101.73)	177–730	218	501.03 (128.25)	136–887
NDW	151	186.03 (29.88)	96–275	218	184.04 (34.85)	77–274
D	156	81.54 (15.62)	27.71–127.20	224	80.04 (15.10)	39.07–128.74
TNC	151	35.19 (15.47)	6–85	218	34.79 (19.99)	3–109
DSS	128	9.68 (1.49)	6.40–13.34	192	9.82 (1.59)	5.74–15.83

Note. MZ = monozygotic twins; DZ = dizygotic twins; MLU-C = mean length of C-unit; NTW = number of total words; NDW = number of different root words used within 100 complete and intelligible child C-units; D = an index of lexical diversity calculated via Vocd (a program available through the Child Language Data Exchange System; http://childes.psy.cmu.edu) using random subsamples of all intelligible root words (McKee, Malvern, & Richards, 2000); TNC = total number of conjunctions used within 100 complete and intelligible child C-units; DSS = developmental sentence score (Lee, 1974), a measure of morphosyntactic complexity based on all complete and intelligible child C-units that contain both a subject and predicate.

Table 2

Partial correlation matrix of child language sample measures, controlling for child age.

Child language variable	1	2	3	4	5	6	7	8
1. MLU-C	—							
2. NTW	.885[a]	—						
3. NDW	.795[a]	.877[a]	—					
4. Measure D	.072	.105	.368[a]	—				
5. TNC	.740[a]	.747[a]	.640[a]	-.185[a]	—			
6. DSS	.645[a]	.561[a]	.534[a]	.127	.471[a]	—		
7. BNT	.181[a]	.162	.249[a]	.284[a]	.046	.162	—	
8. SB-Vocab	.231[a]	.219[a]	.268[a]	.130	.182[a]	.204[a]	.476[a]	—

Note. BNT = Boston Naming Test (Goodglass & Kaplan, 2001); SB-Vocab = Vocabulary subtest from the Stanford–Binet Intelligence Test (Thorndike, Hagen, & Sattler, 1986).

[a]Notes statistical significance at an alpha of .002.

Table 3

Intraclass correlations for monozygotic and dizygotic twins and univariate estimates of genetic and environmental effects on child language sample measures with effects of age and sex removed.

Child language variable	MZr	DZr	h^2 (CI)	c^2 (CI)	e^2 (CI)
MLU-C	.491[a]	.358[a]	.48 (.05–.79)	.09 (.00–.42)	.43 (.31–.61)
NTW	.576[a]	.342[a]	.67 (.32–.89)	.00 (.00–.00)	.33 (.25–.48)
NDW	.531[a]	.301[a]	.61 (.23–.82)	.00 (.00–.00)	.39 (.29–.55)
D	.351[a]	.134	.37 (.00–.56)	.00 (.00–.00)	.63 (.48–.82)
TNC	.384[a]	.268[a]	.49 (.00–.71)	.00 (.00–.00)	.51 (.38–.74)
DSS	.420[a]	.186	.44 (.00–.65)	.00 (.00–.00)	.56 (.40–.76)
SB-Vocab	.467[a]	.218	.45 (.00–.64)	.00 (.00–.00)	.55 (.41–.73)
BNT	.843[a]	.653[a]	.43 (.25–.66)	.44 (.20–.69)	.13 (.10–.18)

Note. CI = confidence interval; h^2 = heritability; c^2 = shared environmental effects; e^2 = combination of nonshared environment and error.

[a] Denotes correlations that are significant at an alpha of .008.

Table 4

Standardized path estimate coefficients from the bivariate analyses.

Variable	Genetic		Shared environment		Nonshared environment	
	A₁	A₂	C₁	C₂	E₁	E₂
	A_1	A_2	C_1	C_2	E_1	E_2
Formal factor[a]	$.67^b$		$.74^b$		.04	
Conversational factor[c]	$.31^b$	$.78^b$	.00	.00	.55	.00

[a] Includes the Boston Naming Test (Goodglass & Kaplan, 2001) and the Vocabulary subtest from the Stanford–Binet Intelligence Test (Thorndike, Hagen, & Sattler, 1986).

[b] Denotes correlations that are significant at an alpha of .05.

[c] Includes mean length of C-unit, number of total words, number of different words, measure D, total number of conjunctions, and developmental sentence scores (Lee, 1974).