

Running head: Infant Language Profiles

Early Language Profiles in Infants at High-Risk for Autism Spectrum Disorders

Kristelle Hudry¹, Susie Chandler², Rachael Bedford², Greg Pasco³, Teodora Gliga⁴, Mayada Elsabbagh^{4,5}, Mark H. Johnson⁴, Tony Charman³ and the BASIS Team

¹Olga Tennison Autism Research Centre, School of Psychological Science, La Trobe University, Australia; ²Centre for Research in Autism and Education, Department of Psychology and Human Development, Institute of Education, University of London; ³Department of Psychology, Institute of Psychiatry, King's College London; ⁴Centre for Brain and Cognitive Development, Birkbeck, University of London; ⁵Department of Psychiatry, McGill University.

Abstract

Many preschoolers with autism spectrum disorder (ASD) present relative lack of receptive advantage over concurrent expressive language. Such profile emergence was investigated longitudinally in 54 infants at high-risk for ASD and 50 low-risk controls, with three language measures taken across four visits (around 7, 14, 24, 38 months). High-risk infants presented three outcome subgroups: ASD, other atypicality, and typical development. Reduced receptive vocabulary advantage was observed in high-risk infants by 14 months, but was maintained to 24 months only in ASD/other atypicality outcome subgroups while typically-developing high-risk infants regained a more normative profile. Few group differences appeared on a direct assessment of language and parent-reported functional communication. Processes of early development toward ASD outcome and in intermediate phenotypes are discussed.

Key Words: Autism Spectrum Disorder, Broader Autism Phenotype, High-Risk Siblings, Receptive Language, Expressive Language, Language Profiles

Corresponding author: Kristelle Hudry, Olga Tennison Autism Research Centre, School of Psychological Science, La Trobe University, Bundoora, VIC 3086, Australia. Phone: +61 (0) 3 9479 5649; Fax: +61 (0) 3 9479 1956; k.hudry@latrobe.edu.au.

Early Language Profiles in Infants at High-Risk for Autism Spectrum Disorders

Substantive language heterogeneity presents within autism spectrum disorders (ASD). While much remains to be understood regarding language processes in these conditions (see Boucher 2012, for a recent review), a profile of relative impairment in receptive skills alongside better-spared expressive abilities appears commonly as a group-level phenomenon (Charman, Drew, Baird and Baird 2003; Ellis Weismer, Lord and Esler 2010; Luyster, Lopez and Lord 2007), holding true for many, but not all, individuals (Hudry et al. 2010; Volden et al. 2011). This contrasts with the normative profile whereby raw receptive language typically exceeds concurrent expressive skills.

At all stages of language acquisition, individuals usually acquire skills first in the receptive domain, with production lagging somewhat behind (see Bates, Dale and Thal 1995). Fenson et al. (1994) present comprehensive vocabulary growth data which demonstrate that single-word comprehension grows rapidly from around 9 months (m) of age. Production follows more slowly, beginning only around 12m and then increasing rapidly later in the second year. Similarly, young infants show early understanding of short phrases and sentences but do not begin to produce these until later, around 18 to 20m (Fenson et al., 1994). Grammatical development follows a similarly staggered pattern (Bates et al. 1995), and while important individual differences in the rate of language acquisition are observable among typically-developing children (e.g., Huttenlocher, Haight, Bryk, Seltzer and Lyons, 1991), the profile of receptive advantage over concurrent expressive skills is consistent.

While language is affected in many developmental conditions, profiles usually retain the normative balance of raw receptive advantage over expressive skills (Ellis Weismer et al. 2010; Laws and Bishop 2003). The contrasting profile often seen in ASD, whereby raw receptive skills fail to show clear advantage over expressive skills, is not a newly-identified phenomenon (i.e., see Bartak, Rutter and Cox 1975). However, attention has recently returned

to this topic, as researchers strive to better understand the extent and underpinnings of this profile (e.g., Charman et al. 2003; Ellis Weismer et al. 2010; Hudry et al. 2010; Luyster et al. 2007; Volden et al. 2011).

A commonly-used parent-report measure, the MacArthur-Bates Communication Development Inventories (MCDI; Fenson et al. 1992), yields meaningful raw counts of receptive and expressive vocabulary (i.e., the numbers of words a child understands and produces). Raw receptive counts *necessarily* exceed raw expressive counts, as children cannot meaningfully use words they do not yet understand. As noted above, Fenson et al. (1994) have shown, within a large normative sample, that infants typically present early and persistent receptive advantage over expressive vocabulary on the MCDI, with acquisition and growth in the former domain preceding the latter by several months. Individuals' expressive vocabulary scores can be expressed as a function of their receptive level. Through comparison with Fenson et al.'s normative data, three independent samples of toddlers/young children with ASD have now been shown to present unusually close alignment across these domains, signifying lack of receptive *advantage* over expressive skills (Charman et al. 2003; Hudry et al. 2010; Luyster et al. 2007).

Various other standardised language/communication measures exist for young children, including direct-assessment tools (e.g., Mullen Scales of Early Learning [MSEL], Mullen 1995; the Preschool Language Scales [PLS], Zimmerman, Steiner and Pond 2002), and parent-report measures (e.g., Vineland Adaptive Behavior Scales [VABS] domain measuring functional communication; Sparrow, Cicchetti and Balla 2005). While assessing different aspects of language, common across such measures is the process of metric standardisation on the basis of data from large, normative samples. Derived Standard Scores (SS) are often of limited use for samples of children with ASD, given frequent floor-level performance. Age-Equivalence (AE) scores are more interpretable but less robust, psychometrically. Given instrument standardisation procedures, close alignment should be

expected across receptive and expressive AEs (as a normative raw advantage in receptive skills should become standardised so as to present balanced receptive and expressive AEs). However, various factors make this an imperfect assumption, including sampling procedures for obtaining the norming group, data interpolation/smoothing procedures undertaken to produce normative score tables, and the substantial differences in performance between individuals resulting in large standard deviations/errors at various age-groupings (Murphy and Davidshofer 1998). Correspondence is also expected across obtained AE scores and a given individual's chronological age (CA), given that the former are derived on the basis of the median CA at which children in the norming group obtained the given score (Murphy and Davidshofer 1998).

In various developmental conditions, including ASD, language AEs frequently present well below CA-expectations (e.g., Ellis Weismer et al. 2010; Laws and Bishop 2003). In conditions such as Down syndrome and specific language impairment, relative profiles of receptive and expressive AEs are usually observed to remain closely-aligned or, if different, to demonstrate heightened receptive compared to expressive AE. However, as detailed below, in the case of children with ASD, data regarding language AE profiles has been highly variable, with scores based on different measures yielding different results, even within one sample.

Broadly speaking, measures of a child's ability to actively demonstrate language comprehension and to produce communication/speech in response to direct examiner bids (e.g., MSEL, PLS, etc.) have consistently shown profiles favouring *expressive* over concurrent receptive AE scores (Ellis Weismer et al. 2010; Hudry et al. 2010; Luyster, Kadlec, Carter and Tager-Flusberg 2008; Volden et al. 2011). By contrast, results based on VABS AE scores, measuring functional comprehension and use of adaptive language/communication (via parent-report and based on every-day observation of the child), have presented highly inconsistently. Ellis Weismer et al. (2010) and Luyster et al. (2008)

reported balanced VABS AE scores, or profiles favouring receptive over expressive AEs, in their samples of toddlers with ‘broad-spectrum’ diagnoses. By contrast, Hudry et al. (2010) reported profiles favouring expressive over receptive AE scores in their sample of preschoolers with core autism.

To date, dedicated investigation of language profiles in ASD has remained cross-sectional. However, a recent developmental surveillance study presented relevant longitudinal data. Barbaro and Dissanayake (2012) examined the MSEL AE scores of infants drawn from a large, community-based sample. While all infants presented early signs of ASD, by 24m, two subgroups were differentiated: those with ASD and with broader developmental/language delays (DD/LD). Receptive language was the domain of most pronounced delay for the infants with ASD outcome, remaining a persistent area of weakness across assessments undertaken at 12, 18, and 24m. While infants with DD/LD outcome also presented early receptive impairments relative to concurrent expressive AEs at 12 and 18m, by 24m the receptive and expressive AEs of this group had become well-aligned. This suggests that while early receptive impairment may signal broadly delayed development, *sustained* lack of receptive advantage may point to emerging ASD. Dedicated research has yet to explore this possibility, and the high-risk sibling design provides just such an opportunity.

Infants at familial high-risk for ASD (hereafter, High-Risk; HR) can be identified on the basis of having a diagnosed older sibling, and can be monitored from early in life through to the point of own possible diagnosis in toddlerhood (e.g., see Zwaigenbaum et al. 2007). This method affords three key benefits to the investigation of early language profiles in ASD. First, it permits repeated assessments to be conducted with children developing ASD, beginning earlier in infancy than otherwise possible/practicable. Second, the inclusion of low-risk controls (with no familial ASD; hereafter Low-Risk; LR) permits direct between-group comparisons rather than reliance on indirect comparison with tabulated norms (see Ellis Weismer et al. 2010 and Luyster et al. 2007 for a discussion). Third, whilst yielding relatively

high saturation of ASD cases, the design also permits evaluation of the broader autism phenotype (BAP; Bolton et al. 1994), characterising features in HR individuals who do *not* go on to ASD. A number of studies have investigated the early language skills of infants developing ASD using the HR design, noting early and continuing receptive and expressive delays in this group (e.g., Landa and Garrett-Mayer 2006; Mitchell et al. 2006; Yirmiya, Gamliel, Shaked and Sigman 2007). Early language delays have presented equally for many HR infants who do *not* proceed toward ASD outcome, indicating potentially important intermediate phenotypes (see Elsabbagh and Johnson 2010). No study has yet, however, specifically examined *profiles* of language ability in HR infants, tracking the normative developmental advantage of receptive over expressive skills in early life.

The current study employed a prospective HR sibling design, presenting the first longitudinal evaluation of language profiles in the context of emerging ASD. We sought to explore patterns in the relative development of receptive and expressive skills across subgroups of infants at HR for ASD, separating three outcome subgroups: individuals with ASD (HR-ASD), other atypicality (HR-Atypical), and typical outcome (HR-Typical). A LR comparison group was expected to display the normative pattern of early-emerging and persistent raw advantage of receptive over expressive skills. We sought to observe whether relative lack of such receptive advantage might emerge early in infants later diagnosed with ASD. We were also interested to examine the specificity of such an early atypical profile to this outcome subgroup, versus more general applicability within intermediate phenotypes of heightened ASD risk.

Given variability in the facets assessed and metrics produced by different language measures, different patterns of result were expected across these. From previous research using the MCDI – a parent-report measure of vocabulary – LR infants were expected to present early and persistent profiles of raw receptive advantage over expressive vocabulary (i.e., large receptive-expressive difference scores). By contrast, HR-ASD infants were

expected to show more closely-aligned receptive and expressive vocabularies (i.e., smaller difference scores). The MSEL – a direct measure of child ability to show understanding of language and to produce appropriate communication/verbalization in response to bids – yields AE scores derived from a large, normative sample. As such, LR infants were expected to show well-matched scores across receptive and expressive domains (i.e., near-zero AE differences), while a profile favouring expressive AE was expected for HR-ASD infants. Also yielding meaningful AEs derived from a large, normative sample, the VABS presents a parent-report measure of everyday functional comprehension and adaptive use of language/communication skills. Again, we expected LR infants to show well-matched receptive and expressive subscale scores, while confident prediction regarding HR-ASD infants was not possible given the inconsistent past results from this measure. Given recent data showing early but transient receptive impairments in infants developing broader developmental/language delays (rather than ASD), we considered it plausible that those HR infants without ASD outcome might present some early profile atypicality, before regaining a more normative receptive advantage.

Methods

Participants

Data from 54 HR and 50 LR infants were made available through the British Autism Study of Infant Siblings (BASIS, www.basisnetwork.org; NHS NRES London REC 08/H0718/76). Each HR infant had an older sibling with a community clinical ASD diagnosis (hereafter, *proband*), confirmed on the basis of information in the Development and Wellbeing Assessment (DAWBA; Goodman, Ford and Richards 2000) and the Social Communication Questionnaire (SCQ; Rutter, Bailey and Lord 2003) by expert clinicians on our team (TC, PB). Most probands met ASD criteria on both measures ($n=44$). While a small number scored below threshold on the SCQ ($n=4$), no exclusions were made due to meeting the DAWBA threshold and expert opinion. For two probands, data were only available on one

measure, and for four probands, neither measure was available (aside from parent-confirmed local clinical diagnosis). Parent-reported family medical histories were examined for significant conditions in the proband or extended family members (e.g., Fragile X syndrome, tuberous sclerosis, condition requiring institutional care) with no such exclusions deemed necessary.

LR controls were full-term infants (gestational ages 38 to 42 weeks) recruited from a volunteer database at the Birkbeck Centre for Brain and Cognitive Development. Medical history review confirmed lack of ASD within first-degree relatives. All LR infants had at least one older sibling (in three cases, only half-siblings). The SCQ was used to confirm absence of ASD in these older siblings, with no child scoring above instrument cut-off ($n=1$ case missing data).

Procedure and Measures

Infants attended up to four visits, each conducted across one or two testing days, around the following mean ages: 7m (range 6-10m), 14m (11-18m), 24m (21-27m), and 38m (32-53m). These are referred to as Visits 1 through 4, respectively. Table 1 presents basic characterisation of the sample, at each visit. Over the course of the study, minimising attrition was considered more important than maintaining strictly-defined age-ranges at each visit (particularly at Visit 4, when attrition would result in missing diagnostic outcome). Groups remained well-matched on mean CA at each visit (all $ps>.38$) and on individuals' inter-visit intervals (all $ps>.32$; Visit 1 to 2 interval, $M=6.4m$, $SD=.93$; Visit 2 to 3, $M=10.2m$, $SD=1.57$; Visit 3 to 4, $M=14.0m$, $SD=3.2$). Non-verbal (NV) ability, the average of MSEL Visual Reception and Fine Motor t -scores (i.e., population $M=50$, $SD=10$), was consistently higher in LR than HR infants.

[Table 1 around here]

Language measures. Parents completed the MCDI (Fenson et al. 1992) at each visit, with minor adaptations for British English. The Words and Gestures (WG) form, used at

Visits 1 and 2, includes ‘understands’ and ‘understands and says’ columns permitting the computation of raw receptive and expressive vocabularies. The Words and Sentences (WS) form, used at Visits 3 and 4, includes only an ‘understands and says’ column. However, parents were asked also to indicate any words only *understood* by the child, permitting assessment of both receptive and expressive vocabulary counts at each visit. This therefore provided repeat parent-report of child knowledge and production of vocabulary, among words commonly-acquired in the early years.

Parents also completed the VABS – 2nd edition (Sparrow et al. 2005), yielding measures of functional communication in receptive and expressive domains. The Parent Rating Form was used at Visits 1 and 2 and at Visit 3 for LR infants. The Survey Interview Form was used at Visit 3 for HR infants and for all participants at Visit 4. Administered by clinical researchers, use of the interview form for HR infants at Visit 3 was considered important to facilitate discussion with parents around any developmental concerns for the child. While the same normative tables are used to derive standardised metrics across both VABS forms, un-facilitated completion by parents might plausibly yield different ratings to those from interview. Though not ideal, lack of consistent use of forms across groups at Visit 3 presents as confounding only for HR-subgroup vs. LR group comparisons, posing no problem for comparisons among HR outcome subgroups.

The MSEL (Mullen 1995) yielded a third set of receptive and expressive scores, via direct assessment of child skills, and was administered by research-reliable assessors. At the toddler visits specifically, administration was by or under close supervision of the same clinical researchers who undertook the diagnostic assessments. Within a direct-assessment context, MSEL scores indexed children’s developing abilities to show understanding of and compliance with language spoken by the examiner, and to produce appropriate communication/language spontaneously and in response to the examiner’s bids/questions (varying across items, as appropriate to the developmental level of the assessed individual).

ASD symptoms and diagnosis. Of the 54 HR infants recruited, 53 were retained to Visit 4 when comprehensive diagnostic assessment was undertaken. Parents completed the Autism Diagnostic Interview – Revised (ADI-R; Lord, Rutter and Le Couteur 1994) and the SCQ (Rutter et al. 2003), and toddlers were assessed with the ADOS-G (Lord et al. 2000; module 2 for 50 toddlers, module 1 for 3 toddlers), with standard algorithms computed. Assessments were conducted without blindness to risk-group status (given the different protocols run with HR and LR toddlers), by or under the close supervision of clinical researchers (i.e., psychologists, speech therapists) with demonstrated research-level reliability.

In determining diagnostic outcome status, four clinical researchers (KH, SC, GP, TC; the first three of whom also conducted the assessments), reviewed information across Visits 3 (including an ADOS-G assessment) and 4 (including both ADOS-G and ADI-R administration). Seventeen toddlers met ICD-10 (World Health Organisation 1993) criteria for ASD (i.e., HR-ASD subgroup). A further 12 toddlers did not meet ASD criteria, but were neither considered typically-developing due either to a) scoring above ADI-R cut-off for autism ($n=1$), b) scoring above ADOS cut-off for ASD ($n=9$), c) scoring greater than 1.5 *SD* below population mean on the MSEL direct-assessment Early Learning Composite ($n=1$), or meeting both of points *b* and *c* ($n=1$). These therefore comprised a HR subgroup presenting other atypicalities (i.e., HR-Atypical), while the remaining 24 toddlers were considered typically-developing (i.e., HR-Typical).

Table 2 provides child outcome characterisation at Visit 4. Comparison of ADOS calibrated severity scores (Gotham, Pickles and Lord 2009) across HR subgroups indicated a significant omnibus effect, $F(2,52)=23.88$, $p<.001$, $\eta^2=.49$, with the HR-Typical subgroup scoring significantly below both HR-Atypical, $t(34)=6.29$, $p<.001$, $d=2.16$, and HR-ASD subgroups, $t(39)=6.43$, $p<.001$, $d=2.06$, which did not differ, $t(27)=.310$, $p=.759$, $d=.11$. Of the 50 LR infants recruited, 48 were retained to Visit 4. One LR infant had attended Visit 1 only, so was excluded from further analysis. Data were retained for the other LR child lost to

Visit 4, who had attended every other visit. While well-matched on Visit 4 CA, $F(3,100)=.85$, $p=.47$, LR and HR outcome subgroups differed on NV ability, $F(3,100)=6.71$, $p<.001$, $\eta^2=.17$. Specifically, HR-ASD toddlers were significantly lower-functioning than LR and HR-Typical toddlers (both $p\leq.006$), with HR-Atypical toddlers falling intermediate.

[Table 2 around here]

Results

A two-staged analytic approach was adopted, with comparisons first made across HR and LR groups to ascertain broader risk-group effects, and then repeated separating out HR-Typical, HR-Atypical and HR-ASD subgroups to evaluate specificity of language profiles to diagnostic outcome. Various data points were missing at each visit, due either to failure to attend a given session or to lack of parental questionnaire completion. MCDI data at Visits 1, 2 and 3, and MSEL data at Visit 3 included 5% or more missing cells (see Appendix Table A1). These appeared to be largely at random with the exception of Visit 1 MCDI scores, where cases with missing data presented lower scores on other language measures at the same visit (e.g., VABS receptive scale). Missing data here may therefore have reflected lack of MCDI completion due to very limited overt child language ability. Missing data were imputed via expectation-maximisation (EM) method (as per Tabachnick and Fidell 2007), with the analyses below presented on the basis of these data. Analyses were also undertaken on the data without imputation, and where slight differences arose, this is noted. Following the presentation of preliminary data handling, the main analyses for each language measure are presented in turn.

Omnibus Analysis of Variance (ANOVA) tests (with Greenhouse-Geisser adjustment applied as necessary) were followed-up with post-hoc multiple comparisons, controlling family-wise error rate via Bonferroni correction. Figures present MCDI difference scores, but MSEL and VABS AE scores are presented for each of receptive and expressive ability separately (rather than presenting AE difference scores), to allow the reader to observe the

overall skill levels of each sub/group, whilst also considering the relative balance across domains (as presented in the text/analysis). Comprehensive tabulation of data is included in Appendix A.

Preliminary Data Handling

Raw MCDI receptive and expressive vocabulary counts were totalled as usual. While clear ceiling effects resulted in the necessary exclusion of Visit 4 data from analysis, for Visits 1 to 3, difference scores were computed by subtracting expressive from receptive counts. Square-root transformation was applied to these data which were highly positively-skewed. This effectively normalised Visit 2 and 3 data, but Visit 1 data remained skewed (i.e., with a modal score of zero, as many infants reportedly neither produced nor yet understood any words). The results of these parametric analyses and of non-parametric analyses conducted on untransformed data were substantively identical, so the former are presented here (including transformation on Visit 1 data). A small number of observed outliers (i.e., data-points falling beyond 2SD of the mean; seven points at Visit 1, four at Visit 2, and three at Visit 3), were replaced with the value one point greater than the highest non-outlier score (as per Tabachnick and Fidell 2007).

MSEL and VABS receptive and expressive AE scores were computed as usual, with receptive-expressive AE difference scores computed as the key metric for analysis. MSEL difference scores were normally-distributed, although some outlier data-points were noted (one low and six high outliers at Visit 2; one low and one high outlier at Visit 4), and handled as above. VABS scores at Visits 3 and 4 were skewed, with square-root transformation effectively normalizing these (and applied also to the non-skewed Visit 1 and 2 VABS data to permit repeated-measures analysis across visits). No outlier data points were apparent.

MCDI Raw Vocabulary Count Difference Scores

A 2 Group (HR, LR) by 3 Visit (mean age: 7, 14, 24m) ANOVA on MCDI vocabulary difference scores revealed significant main effects of Visit, $F(1.49,150.25)=160.30, p<.001,$

$\eta^2=.61$ and Group, $F(1,101)=9.08, p=.003, \eta^2=.08$, and a significant interaction term, $F(1.49,150.25)=3.88, p=.034, \eta^2=.61$ (with the latter non-significant, $p=.059, \eta^2=.04$, when data did not include EM imputation). The intercept term was significantly different from zero, $F(1,101)=872.62, p<.001, \eta^2=.90$, indicating overall receptive advantage over expressive vocabulary (i.e., non-zero difference scores). LR infants demonstrated consistently greater receptive advantage compared to HR infants ($p=.003$). Though both groups showed increasing magnitude of receptive advantage from Visits 1 to 2 (all pair-wise p -values $<.001$), HR infants continued this pattern of increasing magnitude to Visit 3 ($p<.001$), for LR infants, change from Visits 2 to 3 did not reach significance ($p=.066$).

A similar analysis retaining LR infants but separating HR infants by outcome subgroup revealed significant main effects of Visit, $F(1.49,145.74)=111.16, p<.001, \eta^2=.53$ and Subgroup $F(3,98)=4.09, p=.009, \eta^2=.11$, and an interaction term which just failed to reach significance, $F(4.46,145.74)=2.34, p=.051, \eta^2=.07$ (significant at $p=.027, \eta^2=.10$, when analysis did not include EM imputation). Again, the intercept term was significantly different from zero, $F(1,98)=584.69, p<.001, \eta^2=.86$. HR-ASD infants presented significantly reduced receptive vocabulary advantaged compared to LR infants ($p=.012$) with other HR outcome subgroups not significantly different to either of these. While all subgroups displayed increasing magnitude of receptive advantage across Visits (all pair-wise p -values $<.001$), change from Visit 2 to 3 was significant only for HR-Typical infants ($p<.001$), non-significant for LR infants ($p=.066$, as already reported) for HR-Atypical and HR-ASD subgroups (both $p\geq.157$). Figure 1 depicts these effects, presenting group and subgroup mean difference scores (untransformed data).

[Figure 1 about here]

MSEL Direct-Assessment Language AE Difference Scores

A 2 Group (HR, LR) by 4 Visit (mean age: 7, 14, 24, 38m) ANOVA on MSEL AE difference scores revealed a significant main effect of Visit, $F(2.27,229.70)=5.56, p=.003$,

$\eta^2=.05$, but none for Group, $F(1,101)=.002$, $p=.964$, $\eta^2=.00$, nor any interaction term, $F(2.27,229.70)=.23$, $p=.820$, $\eta^2=.00$. While the intercept term was not significantly different from zero, $F(1,101)=1.58$, $p=.212$, $\eta^2=.02$, indicating overall lack of difference between receptive and expressive AEs, at Visit 3, scores evidenced generally greater receptive than expressive AE (i.e., positive mean difference), whereas Visit 1 and 4 scores evidenced generally greater expressive than receptive AE (i.e., negative mean difference; pair-wise p -values $\leq.014$; though again, no significant effects were apparent with statistical control for NV ability).

A similar analysis, retaining the LR group but separating three HR subgroups, also yielded a significant main effect of Visit, $F(2.28,223.87)=3.69$, $p=.021$, $\eta^2=.04$, but none for Subgroup, $F(3,98)=.86$, $p=.464$, $\eta^2=.03$. The interaction term approached significance, $F(6.85,223.87)=2.03$, $p=.054$, $\eta^2=.06$ (reaching significance, $p=.042$, $\eta^2=.06$, when analysis was undertaken without EM imputation). The intercept was not different from zero, $F(1,98)=.72$, $p=.398$, $\eta^2=.01$, again indicating overall balance across receptive and expressive AEs. However, Visit 1, scores evidenced generally greater expressive AEs (i.e., negative mean difference score), consistent across all subgroups, except the HR-Typical group which presented balanced profiles. This pattern at Visit 1 was also significantly different to the more balanced AE profiles observed across all subgroups at Visit 2 ($p=.035$), and marginally different to the balanced AE profiles at Visit 3, although HR-Typical infants here showed heightened receptive AEs and HR-ASD infants showed the reverse. Figure 2 presents group and subgroup MSEL AE scores.

[Figure 2 about here]

VABS Parent-Reported Functional Communication AE Difference Scores

A 2 Group (HR, LR) by 4 Visit (mean age: 7, 14, 24, 38m) ANOVA on VABS AE difference scores revealed a significant main effect of Visit, $F(2.02,204.04)=9.50$, $p<.001$, $\eta^2=.09$ and a trend toward such for Group, $F(1,101)=3.54$, $p=.063$, $\eta^2=.03$ (significant when

analysis was without EM imputation, $p=.035$, $\eta^2=.05$). There was no significant interaction term, $F(2.02,204.04)=.62$, $p=.538$, $\eta^2=.01$. The intercept term differed significantly from zero, $F(1,101)=35.00$, $p<.001$, $\eta^2=.26$, indicating greater overall receptive compared to expressive VABS AE scores. However, this appeared to be driven by Visit 4 data where large receptive-expressive AE discrepancies were apparent (favouring the former), compared to Visits 1 to 3 ($ps\leq.007$) where scores were more balanced. LR infants showed somewhat greater receptive-expressive AE difference scores than did HR infants.

When comparing LR and three HR outcome subgroups, main effects were significant for both Visit, $F(1.97,193.09)=9.49$, $p<.001$, $\eta^2=.09$, and Subgroup, $F(3,98)=4.16$, $p=.008$, $\eta^2=.11$, while the interaction term remained non-significant, $F(5.91,193.09)=1.28$, $p=.271$, $\eta^2=.04$. The intercept term was significantly different from zero, $F(1,98)=13.48$, $p<.001$, $\eta^2=.12$, indicating greater overall receptive compared to expressive AE scores. Again, however, this appeared due to the substantively larger Visit 4 receptive-expressive discrepancies, compared to the more balanced AE scores across Visits 1 through 3 ($ps\leq.014$). Considering subgroups, LR and HR-Typical infants presented somewhat greater AE difference scores than HR-Atypical infants (pairwise, $p=.044$ & $p=.067$, respectively) with HR-ASD infants intermediate. Figure 3 presents group and subgroup VABS receptive and expressive AE scores.

[Figure 3 about here]

Discussion

The current study evaluated early language profiles in infants at familial high-risk for ASD. We sought to prospectively track the emergence of a profile seen in many young children with ASD, characterised by relative *lack* of receptive advantage over concurrent expressive skill. Indices were available across three language measures, taken at multiple visits in infancy and toddlerhood. Our analyses first compared high- and low-risk groups, then separated high-risk infants by diagnostic outcome, as determined after the third birthday.

Language Profiles in Early Development toward ASD

Data for the current sample of high-risk infants with ASD outcome serve to extend, into infancy, the findings from past research sampling young children with established diagnoses. Regarding MCDI vocabulary profiles, the magnitude of receptive advantage over expressive skills shown by infants developing ASD was significantly reduced compared to that of controls. Furthermore, across the second year of life, while low-risk controls tended to experience continued growth in receptive vocabulary advantage, infants developing ASD showed limited enhancement of receptive skills beyond concurrent gains in expressive vocabulary. Ceiling scores unfortunately precluded ongoing evaluation across the third year of life in this cohort. However, the pattern shown here around the second birthday aligns with findings presented elsewhere for toddlers/young children with established ASD diagnoses (i.e., Charman et al. 2003; Hudry et al. 2010; Luyster et al. 2007). Interestingly, our data also show some consistency with patterns observable in the data presented by Mitchell et al. (2006), from a similarly-sized sample of high-risk infants with ASD outcome. Mitchell et al. examined early receptive and expressive language but did not specifically address relative *balance* across these domains. However, comprehensive data tabulation by these authors permits comparison with our own. Around 13m, mean MCDI receptive-expressive difference scores from Mitchell et al. present quite similarly across ASD-outcome and control groups (and intermediate to our own Visit 2 data). However, by 18m, data from Mitchell et al.'s ASD sample align closely with our own 24m data, and present quite differently to data for their sample of controls, showing broad convergence with our own pattern of findings (methodological differences notwithstanding).

Considering profiles obtained from the current MSEL direct assessments of language, a lack of significant intercept term indicates overall alignment across receptive and expressive domains. This is precisely as expected for controls, given procedures of AE metric derivation. Indeed, across all visits, for both controls and high-risk infants with ASD outcome, receptive-

expressive difference scores included confidence intervals spanning zero (with the single exception of Visit 1 scores presenting slight expressive AE advantage in both groups). These results therefore present limited convergence with previous findings using the same and other direct-assessment tools. Previously, MSEL and PLS direct-assessment data have indicated consistently greater expressive compared to receptive AEs in samples of toddlers/young children with established diagnoses (Ellis Weismer et al. 2010; Hudry et al. 2010; Luyster, et al. 2008; Volden et al. 2011). Barbaro and Dissanayake (2012) also demonstrated this on the basis of MSEL direct-assessment scores for their sample of infants developing ASD, showing this profile to persist across the second year of life. Again, tabulated data provided by Mitchell et al. (2006) present a profile slightly favouring expressive over receptive scores on the PLS in 13-month-olds with ASD outcome (although data are SS, rather than AEs). Clear such advantage (on both PLS and MSEL) is then shown at 24m, while controls' scores present as well-balanced across domains, or favour receptive ability.

In the current study, VABS parent-reported functional receptive and expressive communication AEs also presented fairly similarly across controls and infants developing ASD. A significant intercept term, however, indicated a profile generally favouring receptive AE, albeit alongside a main effect of visit. This amounted to near-zero difference scores across Visits 1 to 3 (consistent with assumptions based on AE metric derivation procedures), but a marked receptive AE advantage at Visit 4. At a subgroup level, this Visit 4 effect held for low-risk controls, while for infants with ASD, the mean was in the same direction but with confidence intervals spanning zero. A pattern of well-balanced VABS subscale AEs converges with findings from Ellis Weismer et al. (2010) within a broad-spectrum ASD subgroup (but not within a narrower Autism group). However, this diverges from a previous report by Hudry et al. (2010; showing relative expressive VABS AE advantage) and from Luyster et al. (2008, reporting receptive VABS AE advantage).

Such discrepant VABS parent-reported functional communication findings may stem from sample characterisation differences. Profiles featuring balanced AEs, or profiles favouring receptive score, have arisen in samples of young toddlers presenting broader ASD diagnoses and relatively high-level non-verbal ability (Ellis Weismer et al. 2010, CA=30m, NVAE=20m; Luyster et al. 2008, CA=28m, NVAE=21m), whereas *lack* of VABS receptive advantage was previously reported in a sample of older, lower-functioning children with core autistic disorder (Hudry et al. 2010, CA=44m, NVAE=26m). The current data do converge, therefore, with the former set of findings given this is a relatively high-functioning sample of young toddlers with broad-spectrum ASD diagnoses (i.e., not just core autism), extending past results to suggest that such relative balance in *functional* receptive and expressive communication skills is stable across early infancy and through toddlerhood, on the pathway toward ASD outcome.

Language Profiles in the Context of Intermediate Phenotypes

Given the interest in using high-risk designs to study intermediate phenotypes (Elsabbagh and Johnson 2010), we were equally interested to explore early profiles in those high-risk infants without ASD outcome. BASIS design permitted the delineation of two additional outcome subgroups; 12 infants were considered to present *other* atypicalities (including subthreshold ASD symptoms and/or developmental delay), while 24 were typically-developing. Into the second year of life, all high-risk subgroups presented similar MCDI profiles, quite distinct from those of low-risk controls, suggesting that early atypical receptive vocabulary features within the broader phenotype rather than specifically applicable to ASD outcome. Across the second year, however, clear subgroup differentiation became apparent. Infants with ASD outcome demonstrated minimal growth in receptive advantage across this period. This was true also for high-risk infants developing other atypicalities. However, infants with typical outcome showed clear acceleration in receptive vocabulary

growth over the second year, such that this initially-delayed subgroup presented comparably to low-risk controls by 24m.

As already noted, language profiles based on MSEL direct-assessment presented as quite well-balanced for controls and high-risk infants with ASD outcome, albeit with slight expressive AE advantage at Visit 1. High-risk infants with other atypical outcome presented similarly, with balanced AE profiles sustained across each visit. High-risk infants with typical outcome were initially well-balanced, with fluctuation then observed between Visits 3 (favouring receptive AE) and 4 (presenting clear expressive AE advantage). Again, considering parent-report VABS data for functional communication, profiles for high-risk infants with ASD outcome were consistently well-balanced across visits. However, the patterns observed for the other high-risk subgroups were somewhat different. Infants with typical outcome maintained close correspondence to low-risk controls, with initially well-balanced scores, slight favouring of receptive AE at Visit 3 (despite differential use of VABS forms across high- and low-risk groups at this visit) and significant favouring of receptive AE at Visit 4. Here, however, high-risk infants with other atypical outcome presented differently, more like the ASD-outcome subgroup, with profiles slightly favouring expressive AE at Visit 3, and regaining greater balance at Visit 4.

Elsabbagh and Johnson (2010) proposed that while early perturbations might apply broadly to infants at high-risk, a canalisation process might restore development toward a more typical trajectory and outcome within a majority subgroup. This is consistent with the current MCDI data, whereby an early lag in relative receptive advantage was observable across high-risk subgroups, but maintained only in some individuals (i.e., those with ASD diagnostic outcome or presenting subthreshold symptoms and/or other atypicality) while the majority ‘caught up’ to a more normative profile. The picture presented by our MSEL direct-assessment and VABS parent-reported functional communication AE data is less clear, however, and no existing research has yet provided VABS data for similar prospective

language profiling, in the context of ASD-risk. Mitchell et al.'s (2006) tabulated MSEL and PLS direct-assessment data (SS reported for 12- and 24m visits) suggest well-balanced skills, or profiles favouring receptive SS, within their large subgroup of high-risk infants without ASD outcome (with the exception of 12m MSEL scores favouring the expressive domain). Strong comparison with this existing dataset is limited by methodological differences (subgrouping and metric choices).

Data from Barbaro and Dissanayake (2012) do also support the notion that processes of canalisation of early risk apply in early development beyond the context of heritable ASD risk. In their community-based sample of infants presenting early signs of ASD, an outcome-subgroup with developmental/language delay showed early but transient receptive language impairment, relative to concurrent expressive skills. The early receptive impairments of infants with later-confirmed ASD, however, persisted through toddlerhood. Current MSEL direct-assessment data provide little in the way of corroborating evidence for such canalisation. However, similarity in the pattern of our MCDI raw vocabulary data and these MSEL direct-assessment data is interesting, supporting the need for further prospective, longitudinal investigation of early language profiles in subgroups of infants with and without ASD outcome.

Measuring Receptive and Expressive Language

Consideration of the various available tools for language assessment in young children is required, in attempting to reconcile these results. While available measures span parent-report and direct-assessment types, and routinely differentiate receptive and expressive domains, the specific language/communication facets under evaluation differ across measures and also at different points along a given measure's scale. Furthermore, particulars of the derived metrics are also relevant, and interpreting standardised test scores poses clear challenge with children with ASD (i.e., due to floor effects on standardised scores, or lack of

age-appropriate normative comparison groups). Consideration is limited here to the measures employed in the current study, although these comments apply equally to other measures.

The MCDI (Fenson et al. 1992) provides a parent-report of child knowledge and meaningful production of vocabulary, based on an inventory of words commonly acquired in early language development. Clear benefit lies in the fact that raw data (receptive and expressive vocabulary counts) are meaningful in their own right, removing reliance on norm-referenced scores. While still appropriate for use with many young children with ASD (i.e., who present limited vocabularies), a clear limitation is in the restriction of this measure to early language acquisition. MCDI data collected here at Visit 4 were necessarily excluded from analysis due to many children approaching instrument ceiling. Nevertheless, the MCDI presents perhaps the most useful tool for evaluating relative receptive-expressive skills in young children. Receptive vocabularies *necessarily* exceed expressive counts (i.e., children cannot meaningfully use words they do not yet understand). The magnitude of receptive-expressive difference scores is then easily compared across groups, as in the current study.

Vocabulary knowledge is only one feature of language development, however, and many other tools take a more holistic approach. The MSEL (Mullen 1995) is a direct-assessment of child ability to demonstrate understanding of and compliance with the examiner's spoken language and to produce appropriate communicative or verbal responses to the examiner's bids. Early items test attention to sound and voice, production of vocalization and gesture, and recognition and production of familiar words. Later items tap the understanding of action words, concepts (e.g., size) and simple questions alongside the repetition and spontaneous production of longer verbal constructions. The VABS (Sparrow et al. 2005) takes an approach which is broader still, assessing every-day and functional child communication skills use (including language), via parent-report. Early items tap the ability to direct and sustain attention, and to show functional comprehension and use of various communicative signals including words. Later items address comprehension of and

compliance with longer instructions and child use of speech to gain and provide information to others. The emphasis here is on spontaneous and regular use of skills, as opposed to the optimal performance arguably elicited in the context of a direct-assessment.

As is the case for many other ability and achievement tests, raw MSEL and VABS scores are not readily interpretable and so interpretation relies on comparison of an individual's raw score/s with those obtained by members of a large normative sample. Use of AE metrics is often preferred over other, more psychometrically-robust indices (e.g., SS, percentiles, etc.) with ASD samples, given the greater likelihood that these will produce interpretable data distributed in such a way as to permit parametric analysis (i.e., distributions of SS are often skewed due to floor-level performance). AE scores, as reported here, are derived on the basis of mean raw scores obtained by age-level subgroups within the norming sample. As such, comparability across an individual's various domain AEs might be assumed, but derivation and smoothing procedures mean that some cross-domain AE variability is likely (see Murphy and Davidshofer, 1998). Neither the MSEL or VABS manuals provide guidance on interpreting domain difference scores (e.g., regarding statistically-significant or clinically-meaningful discrepancies). Furthermore, discrepancies are likely to have different implications at different child ages; a 3-month domain difference in infancy is unlikely to equate to one of the same magnitude in toddlerhood.

Comparison across target and control groups, as undertaken here, affords some capacity for interpreting AE difference scores. However, this may also account for our substantive lack of significant VABS parent-report and MSEL direct-assessment findings, given large within-group variability/confidence intervals and substantial between-group overlap. Dedicated research is required to determine the extent to which divergent receptive-expressive domain scores remain within 'normal limits' on a given test, and the point at which such becomes statistically and/or clinically significant. While remaining the most broadly applicable and readily-interpretable norm-referenced scores, results based on the AE metrics

of tests like the MSEL and VABS present far greater ambiguity than the meaningful raw data obtainable in the form of MCDI vocabulary counts.

The current, prospectively-collected MCDI data confirm a lack of normative receptive vocabulary acquisition, relative to concurrent expressive acquisition, in high-risk infants with ASD outcome. They also indicate such atypicality to present relatively early, observable around the same time that the first overt behavioural symptoms of ASD usually appear (Ozonoff et al. 2010), with maintenance through toddlerhood. Lack of receptive vocabulary advantage may therefore present another early sign of emerging ASD, but a relatively subtle one which initially lacks specificity to ASD, reflecting the broader phenotype. *Maintenance* of this early atypicality across the second year of life, however, shows greater specificity to later diagnostic outcome and may present an example of the gradual emergence of behavioural symptoms and/or subtle developmental regression in the early course of ASD, as proposed by Ozonoff et al. (2010). The demonstration of differential pathways of relative vocabulary acquisition in the context of heightened ASD-risk supports notions of compounding atypicality along a pathway toward ASD diagnosis and canalisation of early atypicality toward a more normative trajectory in most high-risk infants as proposed by Elsabbagh and Johnson (2010). Notably, vocabulary counts in this latter subgroup fell well below age-expected levels, receptively and expressively (also fairly consistently true across MSEL direct-assessment of language and VABS parent-reported functional communication scores; see Tables A2 and A3) despite close matching to the low-risk group on outcome non-verbal ability level (see Table 2).

The striking acceleration in relative receptive vocabulary advantage across the second year of life highlights the possible advances to be made in our understanding of ASD and its broader phenotype through the investigation of developmental pathways in the context of early risk. Groups employing high-risk sibling designs regularly separate ‘affected’ and ‘unaffected’ outcomes (e.g., see Mitchell et al. 2006). However, the latter plausibly comprises

substantial heterogeneity in developmental trajectories and outcomes, so dedicated attention those high-risk infants who do *not* present ASD at outcome is warranted in illuminating those processes which serve to steer infants away from disorder, toward more normative outcome.

Strengths, Limitations and Future Directions

Strengths of the current study lay in the recruitment and retention of well-characterised high- and low-risk groups, longitudinal assessment on multiple measures, and rigorous outcome characterisation. Limitations include the relatively modest sample size, particularly separating infants by outcome subgroup. While diagnostic stability of the ASD- and typical-outcome groups is likely, this is less certain for individuals within our *other atypicality* subgroup, who presented sub-threshold ASD traits and/or global developmental delay (see Chawarska, Klin, Paul, Macari and Volkmar 2010). This group tended to present language profiles intermediate to those of the outcome subgroups, and future follow-up will permit observation of symptom consolidation or processes of canalisation within this subgroup. Further recruitment (underway) may permit the more specific delineation of outcome subgroups (e.g., high- vs. low-functioning ASD, sub-threshold ASD only, global delay only, etc.) potentially affording further insight into specific developmental language profiles. Given that some recent reports have suggested atypical receptive-expressive language profiles to present more commonly in individuals with ASD and relatively high-level non-verbal functioning (Hudry et al., 2010; Volden et al., 2011), more detailed analysis of subgroups or covariation among skills sets will be important in future studies with larger outcome groups.

Limitations of the current study include the relatively large between-visit intervals, precluding more detailed observation of early language trajectories across the infant and toddlers years. Future replication with more frequent sampling (particularly during the second year of life) will afford better understanding of those points at which pathways begin to diverge toward differential outcomes. Follow-up beyond the third birthday would also afford more ready comparison of these prospective data with existing cross-sectional data from

samples of young children (e.g., Hudry et al. 2010). Additionally, while basic proband screening was undertaken to confirm parent-reported ASD diagnosis, comprehensive characterisation was not conducted so it is plausible that some probands might have failed to meet full ASD criteria, reducing the relative ‘risk’ of some high-risk infants. This is unlikely to be particularly problematic for the current study as interpretation of our results has been primarily by diagnostic outcome of the infants themselves. This classification was made by a team of experienced clinical researchers, using gold-standard instruments and procedures (though not blind to risk-group status), including case review of data collected across the 24- and 38-month visits, including the ADOS-G and ADI-R.

Finally, a broad limitation regards the lack of any clearly-articulated mechanism underpinning a lack of receptive advantage over concurrent expressive skill, in children with ASD. The current results have provided some novel insights, particularly regarding the early generality and later specificity of observed effects to heightened risk- vs. diagnosed outcome, respectively. The lack of corroborated findings across the three measures employed here, however, precludes us from drawing any stronger conclusions about the underpinnings of relative language profiles at this time. Given that clearest evidence comes from the parent-report MCDI, under-reporting of child comprehension and over-reporting of echolalic speech plausibly play a role in these findings. While Houston-Price, Mather and Sakkalou (2007) have indeed shown discrepancies across parent-reported and preferential-looking indices of receptive vocabulary in typical infants, this seems unlikely to account fully for the phenomenon under current investigation, given the differential subgroup patterns observed here and also the similar pattern reported by other groups (but not replicated here) on the basis of direct-assessments (e.g., Barbaro and Dissanayake, 2012; Ellis Weismer et al. 2010; Hudry et al. 2010; Luyster, et al. 2008; Volden et al. 2011).

Social and cognitive processes involved in language acquisition (e.g., social referencing, recognition vs. recall types of memory, etc.) likely have important roles to play

relative receptive-expressive language acquisition, but have yet to be specifically investigated. The current results provide the very general indication that the mechanism at play is disrupted early in development within intermediate phenotypes, maintaining its effects into toddlerhood as ASD symptoms consolidate, but canalizing its early effects within in the majority high-risk subgroup headed toward typical outcome. Attempts to demonstrate how social-cognitive processes typically produce clear receptive advantage over expressive vocabulary skills, but manifest in the distinct profile and trajectory demonstrated here on the course toward ASD, should now begin.

References

- Barbaro, J., & Dissanayake, C. (2012). Developmental profiles of infants and toddlers with autism spectrum disorders identified prospectively in a community-based setting. *Journal of Autism and Developmental Disorders, 42*, 1939-1948.
- Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific developmental receptive language disorder: I. The children. *British Journal of Psychiatry, 126*, 127-145.
- Bates, E., Dale, P., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language*. Oxford, UK: Blackwell.
- Bolton, P., MacDonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., et al. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry, 35*, 877-900.
- Boucher, J. (2012). Research Review: Structural language in autistic spectrum disorder - characteristics and causes. *Journal of Child Psychology and Psychiatry, 53*, 219-233.
- Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language, 30*, 213-236.
- Chawarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2010). A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. *Journal of Child Psychology and Psychiatry, 50*, 1235-1245.
- Ellis Weismer, S., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *Journal of Autism and Developmental Disorders, 40*, 1259-1273.

- Elsabbagh, M., & Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences, 14*, 81-87.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs for the Society for Research in Child Development, 59*, 1-185.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J. P., et al. (1992). *MacArthur Communicative Development Inventories: User's Guide and Technical Manual*. Baltimore, MD: Paul H. Brookes.
- Goodman, R., Ford, T., & Richards, H. (2000). The Development and Wellbeing Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*, 645-655.
- Gotham, K., Pickles, A., & Lord, C. (2009). Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 693-705.
- Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., et al. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language and Communication Disorders, 45*, 681-690.
- Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language, 34*, 701-724.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236-248.

- Landa, R., & Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: A prospective study. *Journal of Child Psychology and Psychiatry*, *47*, 629-638.
- Laws, G., & Bishop, D. V. M. (2003). A comparison of language abilities in adolescents with Down syndrome and children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *46*, 1324-1339.
- Lord, C., Risi, S., Lambrecht, L., Cook Jr., E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule - Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205-233.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview - Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*, 659-685.
- Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*, 1426-1438.
- Luyster, R. J., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for Autism Spectrum Disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language*, *34*, 623-654.
- Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., et al. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Developmental and Behavioral Pediatrics*, *27*, s69-s78.
- Mullen, B. (1995). *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Services.

- Murphy, K. R., & Davidshofer, C. O. (1998). *Psychological testing: Principles and applications* (4 ed.). Upper Saddle River, NJ: Prentice Hall.
- Ozonoff, S., Iosif, A. M., Baquio, F., Cook, I. C., Hill, M. M., Hutman, T., Rogers, S. J., Rozga, A., Sangha, S., Sigman, M., Steinfeld, M. B., & Young, G. S. (2010). A prospective study of the emergence of early behavioural signs of autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*, 256-266.
- Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales: Survey Form* (2 ed.). Circle Pines, MN: American Guidance Service.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson.
- Volden, J., Smith, I. M., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., et al. (2011). Using the Preschool Language Scale, Fourth Edition to characterise language in preschoolers with autism spectrum disorders. *American Journal of Speech Language Pathology*, *20*, 200-208.
- World Health Organisation. (1993). *Mental Disorders: A Glossary and Guide to their Classification in Accordance with the 10th Revision of the International Classification of Disease - Research Diagnostic Criteria: ICD-10*. Geneva: Author.
- Yirmiya, N., Gamliel, I., Shaked, M., & Sigman, M. (2007). Cognitive and verbal abilities of 24- to 36-month-old siblings of children with autism. *Journal of Autism and Developmental Disorders*, *37*, 218-229.
- Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scale, Fourth Edition*. San Antonio, TX: The Psychological Corporation.
- Zwaigenbaum, L., Thurm, A., Stone, W., Baranek, G., Bryson, S., Iverson, J., et al. (2007). Studying the emergence of autism spectrum disorders in high-risk infants:

Methodological and practical issues. *Journal of Autism and Developmental Disorders*,
37, 466-480.

Appendix

Table A1. Number of missing data-points across three language measures and three/four visits, for LR and HR groups, and HR subgroups (with Typical, Atypical, and ASD outcome).

	LR Group	HR Group Combined	HR Outcome Subgroups		
			Typical	Atypical	ASD
Group N	49	54		12	17
MCDI					
Visit 1	3	4	2	1	1
Visit 2	2 + 1 DNA	3 + 1 DNA	3 + 1 DNA		
Visit 3	1 + 2 DNA	6 + 2 DNA	5		1 + 1 DNA
MSEL					
Visit 1					
Visit 2	1 DNA	1 DNA	1 DNA		
Visit 3	3 ^a + 2 DNA	2 DNA			1 DNA
Visit 4	1 DNA	1 DNA			
VABS					
Visit 1	1	1	1		
Visit 2	1 + 1 DNA	2 + 1 DNA	2 + 1 DNA		
Visit 3	2 DNA	2 DNA			1 DNA
Visit 4	2 + 1 DNA	1 DNA			

N.B. DNA = Did not attend

^a One case missing Receptive AE only

Table A2. Mean (and standard deviation) receptive-expressive difference scores (untransformed data) for LR and HR groups, and for HR diagnostic subgroups (with Typical, Atypical, and ASD outcome), across three language measures and three/four visits.

	LR Group	HR Group Combined	HR Outcome Subgroups		
			Typical	Atypical	ASD
MCDI					
Visit 1	9.43 (16.87)	8.20 (16.69)	9.86 (22.20)	4.36 (7.30)	7.69 (12.54)
Visit 2	92.87 (76.78)	45.98 (54.14)	52.35 (69.29)	44.67 (46.78)	41.18 (40.49)
Visit 3	142.3 (130.3)	105.5 (72.82)	137.5 (76.11)	91.17 (51.29)	76.53 (71.16)
MSEL					
Visit 1	-0.96 (1.81)	-0.57 (2.17)	0.29 (1.90)	-1.08 (1.88)	-1.38 (2.45)
Visit 2	-0.04 (2.77)	0.23 (2.94)	0.04 (2.92)	1.25 (3.05)	0.06 (2.73)
Visit 3	0.91 (5.84)	0.96 (4.74)	2.54 (4.64)	1.42 (4.50)	-1.75 (4.04)
Visit 4	-1.10 (6.06)	-1.55 (6.42)	-3.21 (5.14)	0.50 (5.71)	-0.65 (8.09)
VABS					
Visit 1	1.50 (2.35)	0.45 (3.90)	1.13 (3.01)	-1.25 (3.17)	0.59 (5.17)
Visit 2	1.60 (3.91)	0.12 (2.97)	-0.57 (2.56)	0.33 (2.50)	0.53 (3.57)
Visit 3	3.09 (9.63)	0.12 (8.35)	3.79 (9.34)	-4.50 (4.95)	-1.94 (6.47)
Visit 4	13.54 (27.89)	11.68 (19.12)	16.04 (18.90)	10.42 (21.50)	6.41 (17.26)

N.B. LR and HR groups administered different VABS versions at Visit 3 (Parent Rating Form and Survey Interview Form, respectively)

Table A3. Mean (and standard deviation) receptive and expressive scores (untransformed data) for LR and HR groups across three measures and three/four visits.

	LR Group		HR Group	
	Receptive	Expressive	Receptive	Expressive
MCDI raw score				
Visit 1	9.6 (17.1)	0.2 (0.62)	9.0 (17.2)	0.8 (3.6)
Visit 2	103.8 (86.6)	11.0 (14.9)	57.5 (65.0)	11.6 (23.7)
Visit 3	467.0 (154.2)	324.7 (187.6)	325.9 (167.8)	220.4 (170.1)
MSEL AE score				
Visit 1	6.8 (1.5)	7.7 (1.9)	5.8 (2.3)	6.4 (1.9)
Visit 2	13.8 (3.2)	13.8 (2.4)	12.8 (3.7)	12.6 (3.4)
Visit 3	28.8 (4.3)	27.9 (5.5)	25.1 (6.3)	24.1 (6.1)
Visit 4	43.6 (7.8)	44.7 (6.9)	38.3 (9.3)	39.8 (9.9)
VABS AE score				
Visit 1	8.7 (3.0)	7.2 (1.9)	6.7 (3.4)	6.2 (2.8)
Visit 2	15.7 (4.8)	14.1 (3.4)	12.1 (4.5)	12.0 (5.2)
Visit 3	32.3 (11.6)	29.2 (6.9)	26.5 (10.1)	26.4 (6.1)
Visit 4	66.8 (35.2)	51.0 (13.1)	51.9 (26.2)	40.3 (12.4)

N.B. LR and HR groups administered different VABS versions at Visit 3 (Parent Rating Form and Survey Interview Form, respectively)

Table A4. Mean (and standard deviation) receptive and expressive domain scores (untransformed data) for HR diagnostic subgroups (with Typical, Atypical, and ASD outcome) across three language measures and three/four visits.

	HR-Typical		HR-Atypical		HR-ASD	
	Receptive	Expressive	Receptive	Expressive	Receptive	Expressive
MCDI raw score						
Visit 1	11.0 (22.4)	1.1 (4.9)	4.9 (7.8)	0.6 (1.2)	8.3 (14.0)	0.6 (2.5)
Visit 2	64.0 (76.1)	11.7 (16.6)	55.4 (59.5)	10.8 (15.7)	53.9 (58.7)	12.7 (34.9)
Visit 3	340 (119)	203 (143)	293 (174)	202 (162)	334 (217)	258 (210)
MSEL AE score						
Visit 1	6.2 (2.0)	5.9 (1.4)	5.6 (1.9)	6.7 (1.9)	5.4 (3.1)	6.9 (2.5)
Visit 2	13.7 (3.9)	13.7 (3.4)	13.1 (4.1)	11.8 (3.2)	11.7 (2.9)	11.6 (3.4)
Visit 3	26.4 (6.3)	23.9 (5.5)	25.6 (3.2)	24.2 (5.1)	22.6 (7.7)	24.4 (7.9)
Visit 4	40.6 (5.9)	43.8 (7.9)	36.7 (5.1)	36.2 (8.3)	36.0 (14.1)	36.7 (11.8)
VABS AE score						
Visit 1	7.3 (3.3)	6.2 (2.6)	5.2 (3.6)	6.4 (2.6)	6.7 (3.3)	6.1 (3.2)
Visit 2	13.2 (3.8)	13.8 (4.4)	11.2 (5.7)	10.8 (5.6)	11.4 (4.6)	10.9 (5.4)
Visit 3	29.7 (10.2)	25.9 (5.3)	22.8 (6.5)	27.3 (4.1)	24.6 (11.1)	26.5 (8.3)
Visit 4	59.2 (24.1)	43.2 (12.5)	50.3 (28.1)	39.9 (11.2)	42.8 (26.0)	36.4 (12.5)

Figure Captions

Figure 1. Mean raw (untransformed) receptive-expressive difference scores (MacArthur-Bates Communication Development Inventories; Fenson et al. 1992) for two risk-groups, and for four outcome subgroups, across three visits. Error bars represent 2x Standard Error.

Figure 2. Mean direct-assessment receptive and expressive language age-equivalence scores (Mullen Scales of Early Learning; Mullen 1995) for a) two risk-groups and b) four outcome subgroups, across each of four visits. Error bars represent 2x Standard Error.

Figure 3. Mean (untransformed) parent-reported functional receptive and expressive communication age-equivalence scores (Vineland Adaptive Behaviour Scales – 2nd edition; Sparrow et al. 2005) for a) two risk-groups and b) four outcome subgroups, across each of four visits. Error bars represent 2x Standard Error.

Figure 1 Top

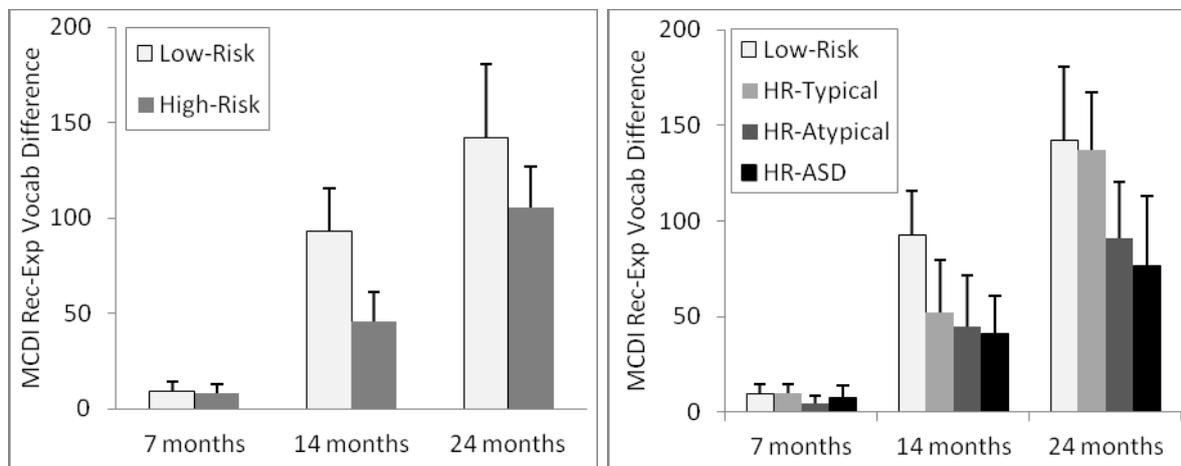


Figure 2 Top

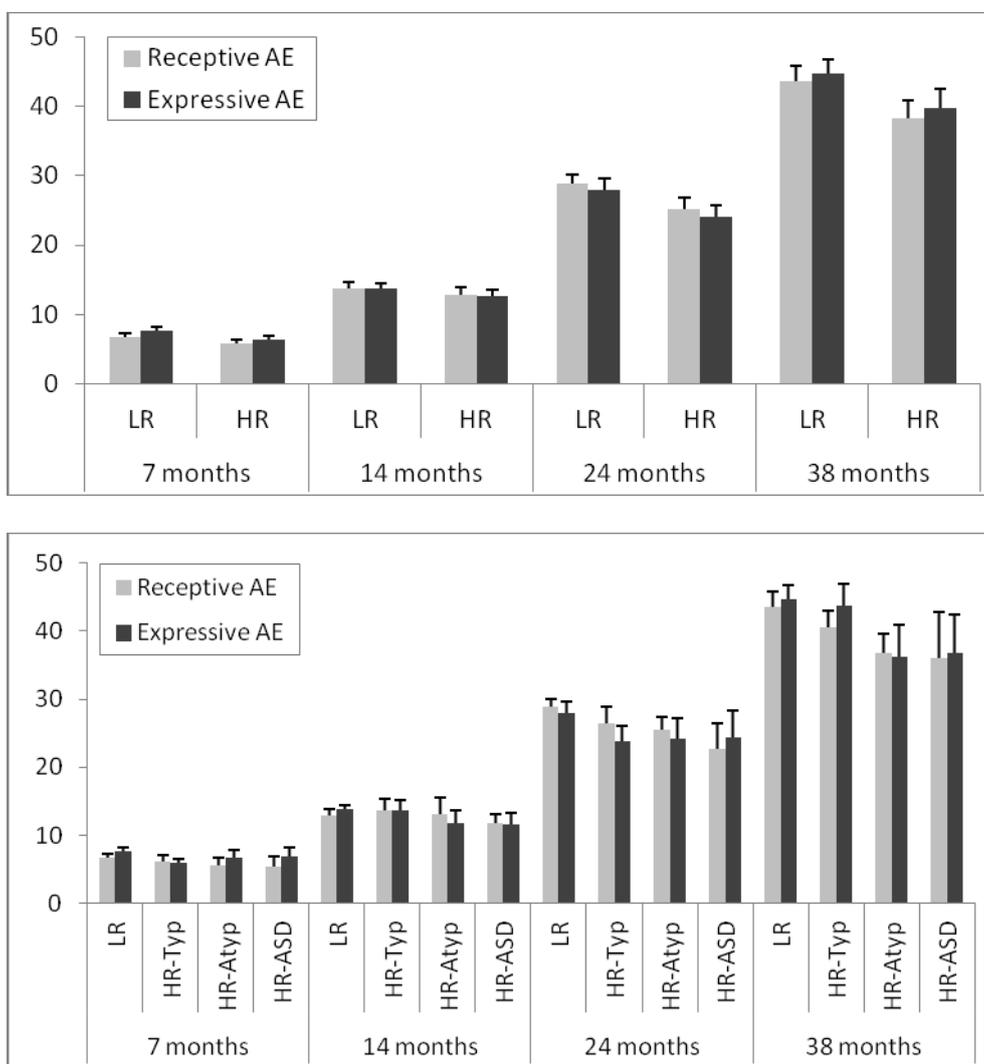


Figure 3 Top

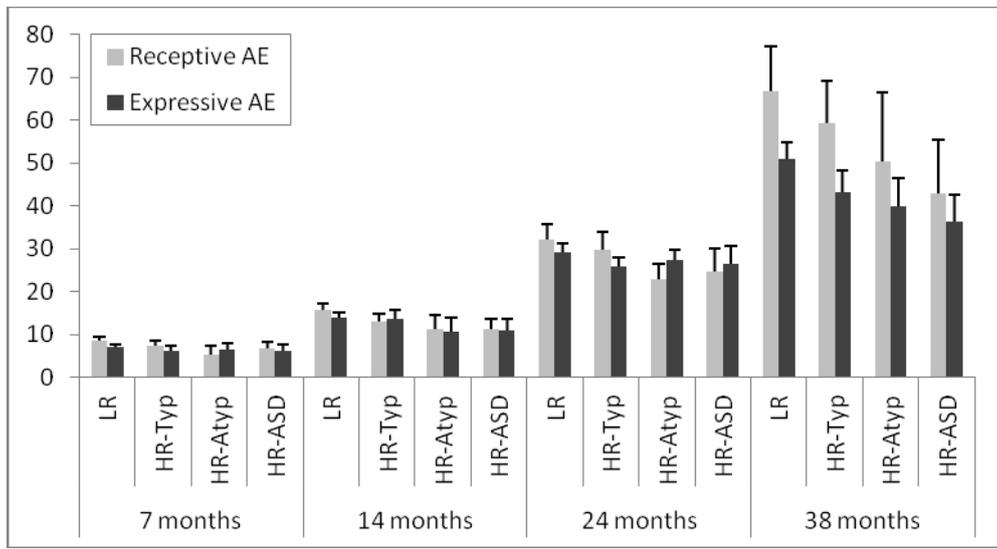
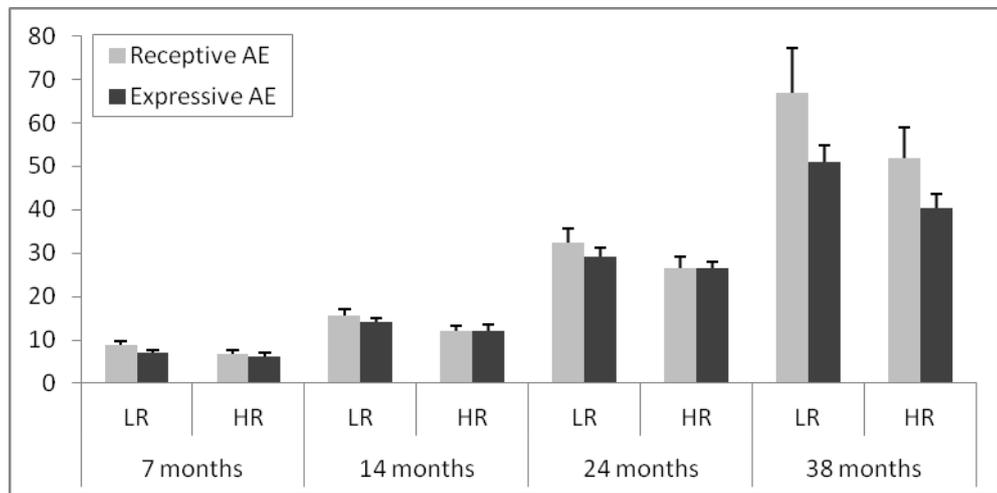


Table 1. Characterisation of infants at high-risk for ASD and low-risk controls, at four visits.

Visit	High-Risk (39% male)			Low-Risk (42% male)			NV group comparison
	N	CA ^a	NV ^b	N	CA ^a	NV ^b	
1	54	7.4 (1.2)	51.5 (8.4)	50	7.4 (1.3)	56.2 (7.1)	$t(101)=3.04, p=.003, d=.42$
2	53	13.7 (1.6)	53.1 (10.3)	48	14.0 (1.3)	58.5 (8.3)	$t(98)=2.87, p=.005, d=.39$
3	52	23.9 (1.2)	51.6 (9.7)	47	23.7 (0.7)	56.5 (9.3)	$t(95)=2.55, p=.012, d=.35$
4	53	37.7 (3.0)	52.0 (13.5)	48	38.2 (3.1)	57.8 (9.9)	$t(95)=2.47, p=.015, d=.34$

^a Group mean (and standard deviation) Chronological Age in months.

^b Group mean (and standard deviation) Non-Verbal ability T-score (based on average of Visual Reception and Fine Motor MSEL domains; Mullen 1995).

Table 2. Detailed characterisation of HR subgroups and LR controls at outcome Visit 4.

	High-Risk			Low-Risk
	ASD (N=17)	Atypical (N=12)	Typical (N=24)	(N=48)
% Male	65%	25%	29%	42%
CA ^a	37.8 (2.1)	36.7 (1.8)	38.1 (3.9)	38.2 (3.0)
MSEL NV ^b	44.00 (16.2)	53.17 (12.1)	57.08 (9.3)	57.79 (9.9)
SCQ ^c	10.9 (7.9)	5.3 (8.9)	3.8 (3.5)	3.0 (2.4)
ADI-R ^c Social Interaction	9.75 (5.5)	3.42 (4.9)	1.63 (1.7)	
Communication	8.44 (5.1)	3.58 (5.5)	2.21 (1.8)	
Repetitive Behaviour	3.63 (2.2)	1.08 (1.3)	0.50 (0.9)	
ADOS Severity	6.12 (2.6)	5.83 (2.2)	2.25 (1.2)	

^a Chronological age (in months)

^b Mullen Scales of Early Learning (Mullen 1995) average of Non-Verbal *t*-scores

^c Social Communication Questionnaire (Rutter et al. 2003); *N.B.* LR group scores ranged from 0 and 7 with the exception of one child who scored 12.

^d Autism Diagnostic Interview – Revised (Lord et al. 1994)

^e Autism Diagnostic Observation Schedule calibrated severity score (Gotham et al. 2009)

Author Note

Kristelle Hudry; Olga Tennison Autism Research Centre, School of Psychological Science, La Trobe University, Melbourne, Australia. Susie Chandler, Rachael Bedford, Greg Pasco, and Tony Charman; Department of Psychology, Institute of Psychiatry, King's College London, United Kingdom. Teodora Gliga, Mayada Elsabbagh, and Mark Johnson; Centre for Brain and Cognitive Development, Birkbeck, University of London, United Kingdom. Mayada Elsabbagh; Department of Psychiatry, McGill University, Montreal, Canada. The BASIS Team in alphabetical order: Simon Baron-Cohen, Patrick Bolton, Janice Fernandes, Holly Garwood, Leslie Tucker, and Agnes Volein.

We would like to thank the BASIS families for their participation in this study, and Carina de Klerk, Kathryn Frame, Besterah Kultu, and Emma Tait for their research assistance. Many thanks also to three anonymous reviewers for their comments on a previous version of this manuscript. The research is supported by The UK Medical Research Council (G0701484) to M. H. Johnson, the BASIS funding consortium led by Autistica (www.basisnetwork.org) and a grant from Autism Speaks (1292). Further support for some authors is from COST action BM1004.

Correspondence concerning this article should be addressed to Kristelle Hudry, Olga Tennison Autism Research Centre, School of Psychological Science, La Trobe University, Bundoora, VIC 3086, Australia. Phone: +61 (0) 3 9479 5649; k.hudry@latrobe.edu.au.

The authors declare that they have no conflict of interest.