Genes Take Over When the Input Fails: A Twin Study of the Passive

Jennifer B. Ganger[1], Sabrina Dunn[1], Peter Gordon[2]
University of Pittsburgh; Teachers College, Columbia University

That language development is the result of some combination of genetic and environmental forces is, uncontroversial. But rarely are we able to specify genes and environment interact specifically within the actual structures of language itself. The twin method, a technique that uses identical and fraternal twins to estimate the relative contributions of heredity and environment to variation in a trait, can help to pinpoint the relative contributions of these factors in language development. Further, since we know about the structure of language in exquisite detail, it is possible to use the twin method to estimate heritable and environmental effects for particular constructions and then, further, ask how the estimated heritability is affected by the availability of relevant examples in the child's linguistic input.

In the present study, we consider the passive voice as a domain in which these issues may be fleshed out. The acquisition of the passive construction has long been the subject of debate over the influence of biological or environmental factors in determining the rate of acquisition. . In English speaking children, the passive construction emerges relatively late in development, and continues to be the source of errors well into the sixth year of life. This has prompted speculation that crucial grammatical structures that are required to represent passive constructions are maturationally delayed (Borer & Wexler, 1987, 1992; Wexler, 2005). However, passivized verbs are also rare in children's language input, suggesting a simpler account: that children do not have sufficient exposure to the passive construction to learn it fully. In the present study, we capitalize on individual differences in the timing of the emergence of the passive in English, and use the twin method to ask to what extent such variation is heritable (that is, due to genetic differences) and to what extent environmental differences are responsible.

1. The Twin Method

The twin method uses differences between correlations for identical and fraternal twins. Identical, or monozygotic (henceforth MZ) twins share 100% of their genes. Fraternal, or dizygotic (henceforth DZ) share, on average, 50% of their genes, just as non-twin full siblings do. However, both types of twins share equally similar environments (i.e., they are raised in together). Therefore if MZ twins are more highly correlated than DZ twins with respect to some trait, the increased correlation can only stem from their increased genetic correlation. The *heritability* of a trait (h^2), or the extent to which its variation can be accounted for by genetic variation, is equal to twice the difference between MZ and DZ twin correlations ($h^2 = 2*(r_{MZ} – r_{DZ})$. The remaining variance comes from environmental variation, which is divided into two sources: shared and non-shared environment. *Shared* or *common environment* (c^2) represents variation due to all environmental factors that make family members similar. It can be calculated by subtracting the heritability from the MZ correlation (i.e., $c^2 = r_{MZ} – h^2$). *Non-shared* or *unique environment* (u^2), which is variation stemming from environmental sources that make family members different, can be estimated as $1 – r_{MZ}$ (Plomin, DeFries, McClearn, & McGuffin, 2001)

2. The Passive

The acquisition of the passive voice has long been the subject of theoretical and experimental investigation in language development. It is a paradigm example of syntactic rule learning. In English, word order is fundamental in conveying syntactic structure and sentence meaning, and the passive construction capitalizes on a meaning change that emerges there is alternation of word order. The noun phrase that typically carries the role patient/theme/recipient canonically appears as the grammatical object in declarative sentences (e.g., Mary closed THE DOOR) but appears as the grammatical subject in passive sentences (e.g., THE DOOR was closed by Mary). The do-er, or *agent* of the action, is relegated to a *by*-phrase ("by Mary") or dropped altogether (THE DOOR was closed.)

English-speaking children have great difficulty with comprehension of the passive construction up to 4 or 5 years of age. Furthermore, passives of verbs denoting observable actions, the so-called "actional" passives (e.g., *the ball was kicked by John*), are mastered earlier than non-actionals (e.g., *the radio was* *heard by John*) (Gordon & Chafetz, 1990; Maratsos, Fox, Becker, & Chalkley, 1985; Pinker, LeBeaux, & Frost, 1987) These facts have prompted two classes of explanations, environmental/external and genetic/internal.

Passives are both late to be acquired and rare in the linguistic input of children. Gordon and Chafetz (1990) examined the prevalence of actional and non-actional passives in the transcript data from Adam Eve and Sarah in CHILDES (Brown, 1973; MacWhinney & Snow, 1985) and found that, while actional passives are moderately rare, non-actional passives—which children are slower to acquire—are vanishingly rare in the linguistic input to children (although adjectival passives showed less asymmetry in this respect). Gordon & Chafetz (1990) also found that children were consistent when tested twice on test and re-test on the same verbs. The verbs that they were correct on the first test tended to be the same as those they were correct on the re-test. Such consistency provides evidence for verb-by-verb learning of passives. This account is bolstered by experimental studies by Tomasello and colleagues (Brooks & Tomasello, 1999) showing children younger than 3 years do not spontaneously use a nonce verb in the passive voice if it has only been modeled in the active voice. This finding led them to argue that grammatical constructions like the passive are initially learned imitatively on a verb-by-verb basis and that generalizations do not emerge until the child has sufficient experience (starting at around 3;5 in Brooks & Tomasello). An additional piece of evidence that input matters comes from (Demuth, 1989), who reports that children learning Sesotho, are provided with abundant examples of passive sentences in everyday speech, and consequently they regularly use passives from the age of two years. Thus, it seems reasonable to believe that the low frequency of the passives in children's input is responsible not only for the delay in the acquisition of passives in English, but also for differences in the rate that actional and non-actional passives are acquired.

Other theories suggest that there may also be internal processes at work in accounting for the slow pace of acquisition for non-actional passives. Pinker et al. (1987) argue that, cross-linguistically, actional passives, in which the theme/patient/recipient is highly affected, are universally licensed by grammar. On the other hand, passives formed from non-actional verbs are rare across languages and that the ease of acquiring actional passives reflects innate linguistic dispositions, whereas non-actional passives reflect language-specific differences that must be learned. A training study with novel verbs confirmed that pre-schoolers learned novel actional passives (the duck was daxed by the mouse) more easily than non-actional passives (Pinker et al.1987). They suggest that the advantage for actional passives occurs independent of experience with those verbs. Non-actional passives, on the other hand, correspond to semantic generalizations other than affectedness of the patient/theme/recipient (or have less affectedness), and must be licensed by experience. Since non-actional passives are rare in the input, English-speaking children require more time to modify their grammars to include these forms. Thus this theory relies on genetic constraints in the grammar but also allows a role for experience in modifying the grammar.

Another approach to the genetic/internal process account of passive acquisition is Borer & Wexler's *maturation of A-chains* or A-Chain Deficit hypothesis (ACDH) (Borer & Wexler, 1987; 1992), which has been recast as the *Universal Phase Hypothesis* (UPP) in more recent writings (Wexler, 2005) The heart of this approach depends on the notion of the A-chain or "argument chain". An A-chain is the link between the position a patient or theme should be in—i.e., the object position (John kicked THE BALL)—and the position in which it appears with the passive—i.e., the subject position (THE BALL was kicked by John). The chain is illustrated in (1) as an operation that binds the sentence Subject to its trace in Object position.

(1) [THE BALL]$_i$ was kicked t_i by John.

Borer & Wexler suggested that the ability to form an A-chain starts to mature sometime after age 4. A similar account applies to the acquisition of unaccusatives (Babyonyshev, Ganger, Pesetsky, & Wexler, 2001), which also require A-chains (Perlmutter & Postal, 1984), and raising constructions (Wexler, 2005), which require a similar operation at LF. Wexler and his colleagues argue that innate knowledge is like physical development and matures along a genetically-determined timeline. On this view, maturation of the relevant neural tissue, not paucity of input, is responsible for late development. Borer & Wexler, suggest that children initially treat passives formed with action verbs as adjectival (stative) rather than verbal (actional) and this explains differences in performance on actional versus non-actional passives. Actional passives like "the door was closed" are claimed to be treated as stative by children as if they were like the adjectival construction: "the door was stuck" rather than like the clearly verbal construction: "the door was slammed". Crucially, they claim that actional passives are more easily construed as adjectival than are non-actional passives. Since adjectival passives are not derived through movement of arguments, there is no A-chain involved, and hence they would not be subject to the purported maturational constraints that would retard the development of non-actional passives.

These theories leave us with three radically different views of the delay in non-actional passives (and passives generally). The class of hypotheses that rely on external factors would predict that differences among individuals are determined primarily by variation in input. For instance, acquisition of actional passives could be correlated with the amount of input that they receive. On the other hand, since almost no non-actional passives occur in the input, then one might not expect variability in rate of acquisition to be predicted by the impoverished input, which would not provide enough variability to be predictive. In this case, acquisition might only come through generalization mechanisms, either linguistic or cognitive in nature. To the extent that the efficiency of such mechanisms is subject to genetic variability, then one might predict that one would find greater heritability in this case. The class of "internal" hypotheses that point to the default nature of the actional passive cross-linguistically as evidence for its genetically determined status, would place little store on the importance of environmental input in the timing of acquisition for actional passives. On the other hand, because non-actional passives are exceptional, they could only be acquired through experience and should exhibit lower heritability and greater effects of common environment (c^2). The maturational account predicts that heritability should be strong in all cases since genetic contributions would determine the timing of maturation, which occurs for both actional and non-actional passives. However, if children's parsing of actional passives really does involve the ghost adjectival interpretation, then it is unclear whether genetic or environmental variation would account for this development.

3. Methods

3.1 Participants

The sample reported herein contains 69 MZ and 117 DZ pairs of twins with a mean age of 5.09 years (s.d. = .94), recruited between 1999 and 2004 in the Pittsburgh PA area. The age, zygosity, and sex distribution of the sample is given in Table 1. Twins were recruited through Mothers of Twins Clubs, a twin registry at Magee-Womens hospital in Pittsburgh (Strassberg et al., 2002), flyers at pediatricians' offices, and word of mouth. On a scale where 0 denotes less than high school education, 1 denotes high school, 2 denotes some college, 3 denotes college degree, and 4 denotes any post graduate education, the average level of maternal education was 2.83 (sd = .93) and the average level of paternal education was 2.55 (sd = 1.09). Due to unavoidable biases in recruitment, the sample was almost entirely Caucasian (94% Caucasian, 1% Hispanic, 4% non-responders, and less than 1% Pacific Islander).

Table 1: Age, sex, and zygosity distribution of subjects

Age	MZ-F	MZ-M	DZ-F	DZ-M	DZ-OppSex
3;0-3;11	2	4	8	5	6
4;0-4;11	16	12	13	5	13
5;0-5;11	11	11	8	8	25
6;0-6;11	9	4	5	9	12
Sum	69 MZ pairs		117 DZ pairs		

3.2 Materials and Procedure

Comprehension of the passive voice was tested with a variation of a test used by Gordon & Chafetz (1990). Historically, most tests of passive comprehension require the child to act out a test sentence or choose between 2 or more pictures depicting actions to test understanding. Such tests require that test sentences be *reversible*—that is, they require that the subject and object be equally plausible if reversed. Without this constraint, children might use common sense to choose their answer, rather than true comprehension of the passive. However, act-out tasks and picture choice tasks have two flaws. First, they require another layer of processing beyond comprehension. Second, the requirement that the passive sentences be reversible makes it impossible to test non-actional passives, such as *watched, held,* and *forgotten*.

Gordon & Chafetz's method solved this problem in a clever way. Each test item begins by telling a brief story. For example, "John's favorite movie is the Wizard of Oz. It was on TV, so he stayed inside to watch it." The child then hears two questions, both in the passive voice: "Was John watched by the movie?" and "Was the movie watched by John?" (The child is trained in advance that he must answer "yes" to one of the questions and "no" to the other using a 10-item warm-up depicting simple spatial relationships between objects. If the child does not get three in a row correct on this warm-up, he does not continue with the passive test.) The child cannot use the heuristic of plausibility to formulate his response, because "John" and "movie" appear in both sentences. If the child understands the passive voice perfectly, he will answer "no" to the first question and "yes" to the second. If he does not understand the passive and answers randomly, his chance performance will be reflected in his overall score.

The passive battery has 18 items: (*watch, drop, forget, hear, eat, carry, kiss, hold, believe, remember, wash, know, like, see, kick, hate, shake,* and *hug.*) The battery used in the present study is identical to that of Gordon & Chafetz except that the original test contained an equal mix of active sentences, short passive sentences (i.e., no *by*-phrase, as in *the door was closed*) and full passive sentences (i.e., with *by*-phrase), whereas the current version contains only long passives. This was change was made after pilot testing revealed that children's performance on the long passive improved when other sentence types were not intermixed. This change also provided more opportunities for each child to show his comprehension of the passive, which potentially increases variance.

Twins were tested in their own homes, in most cases by two different experimenters to avoid bias on the basis of zygosity. They were separated into two different rooms or areas of the house and completed 3 or 4 tests of language or intellectual development, of which the passive battery was one.

3.3 Analysis

Since some children could not answer all 18 items on the passive comprehension task, each child received a score consisting of the percent of correct responses, with those who answered fewer than 8 items (at least 4 actional and at least 4 non-actional) not included in the analyses. Using the DeFries-Fulker method (DeFries & Fulker, 1985), this score was entered into a regression: $T = b_1C + b_2R + b_3CxR + e$, where T is one twin's score, C is her co-twin's score, R is the genetic relationship between the twins (1.0 for MZs and .5 for DZs), CxR is the product of the co-twin's score and the twins' genetic relationship, and e is leftover variance (error). Since there is no principled way to decide which twin should be T and which C, each twin pair is entered twice (so-called double-entry). (Standard errors and p-values are adjusted by hand to compensate for the change in degrees of freedom.) The coefficient b_1 provides an estimate of common environment and b_3 represents heritability. These quantities are mathematically equivalent to the simpler calculations given above, but using regression allows straightforward calculation of standard errors as well as the possibility of including other regressors, such as age and sex.

4. Results

The mean proportion correct for all items and all children was .63 (s.d. = .21). However, the mean percent correct for actional passives was .75 (s.d. = .24) versus .52 (s.d. = .25) for non-actional passives, a significant difference ($t = 17.32$, $p < .0001$) (see Figure 1). An ANOVA with zygosity, sex, and age as random-effects independent variables was performed. Zygosity was not a significant predictor of passive performance (i.e., MZ and DZ twins did not show mean differences). There was a borderline effect of sex ($p = .043$), with boys performing slightly better than girls. There was, however, a significant effect of age ($F = 22.3$, $p < .001$). Therefore, age was included as an additional independent variable in the genetic analysis below.

For all passives, the first order correlation between MZ twins was .55 (partial correlation with age removed: .45), DZ r = .44 (partial r = .35). For actional passives only, the MZ correlation was .43 (partial r = .35), DZ r = .47 (parial r = .40). For non-actional verbs, MZ r = .53 (partial r = .46), DZ r = .30 (partial r = .28). (See Figure 2).

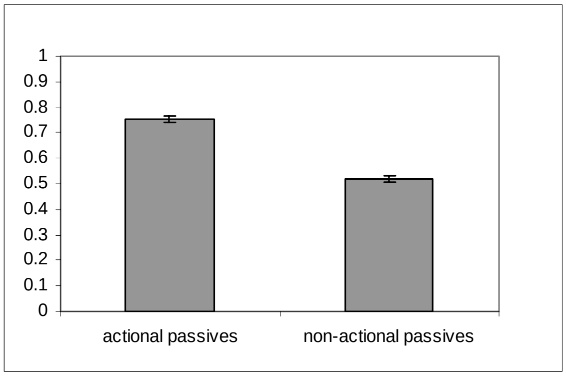

Figure 1: Mean proportion correct (based on co-twin average) for actional versus non-actional passives.

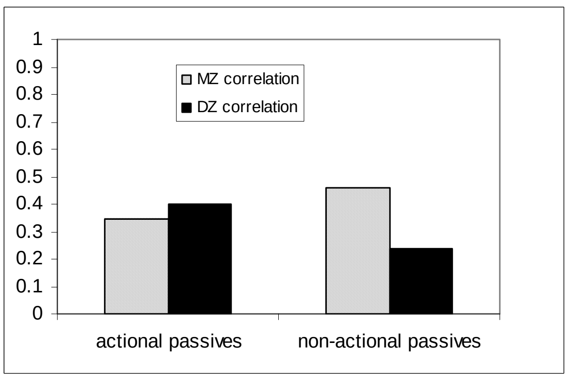

Figure 2: MZ and DZ correlations

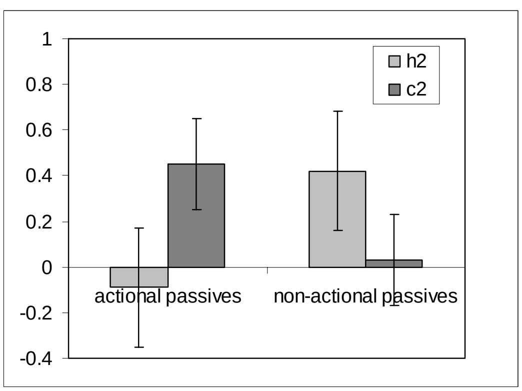

Figure 3: Heritability and shared environment for actional and non-actional passives.

In the behavior genetic analysis, for all passives combined, with age added as a regressor, h^2 = .22 (95% CI: -.30 to .73) and c^2 = .24 (-.14 - .61). Thus, there were small effects of both heritability and shared environment, with a larger role for non-shared environment. Within passive types, we found that h^2 for the 9 actional items was considerably less than for non-actional items: h^2 (actionals) = -.09 (95% CI: -.61 to .43), h^2 (non-actionals) = .42 (95% CI: -.13 to .98) (see Figure 3). (A negative heritability is typically interpreted as zero, since these quantities are components of variance and cannot be negative. This seems reasonably since, as the CI indicates, the actional h^2 point estimate is not reliably different from zero.) Shared environment showed the opposite pattern: For actionals, c^2= .45 (95% CI: .07 to .83), for non-actionals c^2= .03 (95% CI: -.39 to .40). Non-shared environment remained steady at around .55.

5. Discussion

The non-genetic results are quite consistent with previous research: 3- to 6-year-olds are still mastering the passive voice, and they are much more successful with actional passives than with non-actionals. The goal of the genetic analysis is to gain insight into this contrast.

Although the heritability of the passive overall was not high, a striking dissociation emerged when actionals and non-actionals were considered separately. Actionals showed little effect of heritability and large shared environment, while non-actionals showed the opposite profile: substantial heritability and little shared environment. Non-shared environment was moderate to large in all cases. Note that non-shared environment includes not only the idiosyncratic life experiences of twins, but also any random error introduced by the experiment. Such error is inherent in all behavioral studies, but in twin studies it gets dumped into non-shared environment, making this quantity difficult to interpret. Thus, we have not attached much significance to the large non-shared environment in this study.

The main finding of low h^2/high c^2 for actionals and high h h^2/low c^2 for non-actionals allows us to discriminate to some extent among the accounts of the acquisition of passives considered above. It appears to be incompatible with Pinker et al's theory that learning non-actionals is highly dependent on experience, since heritability, not environmental factors, were the main determinant of individual differences.

The finding is at best ambiguous with respect to Borer & Wexler's ACDH/UPP. If Borer & Wexler's suggested mapping between actional passives and adjectival passives (and non-actional passives and verbal passives) is complete, then these results are compatible with their theory. Since non-actionals cannot be interpreted as adjectives, they require A-chains (or they define a defective *v*, in Wexler's recent formulation). Thus, unlike actional passives, which can be interpreted as adjectives without an A-chain, non-actionals require an A-chain and therefore provide the true test of maturation of A-chains. To the extent that non-actionals showed high heritability, the ACDH/UPP account is supported.

However, the mapping between actional and adjectival passives is tenuous. Although such a trend may exist, one must examine the particular verbs used in our test to determine whether the mapping holds. Using both intuitions and tests from Wasow (Wasow, 1977), it appears the mapping is not perfect. Furthermore, Embick (Embick, 2004) suggests the situation may be more complicated, with some adjectival passives (but not all) being formed in the syntax, though differently from the formation of verbal passives. We therefore leave this as a question for future research.

Theories of the importance of input fared better. Gordon & Chafetz (1990) showed that non-actional passives are less frequent in the input than actionals, and it is exactly this class of passives—the non-actionals—that showed higher heritability. Thus, one possibility is that when there is no input available, children must rely on internal mechanisms that are subject primarily to genetic (rather than environmental) variability. These data suggest that there is an interesting interaction between the action of genes and the availability of information in the environment. In the case where information is abundant (actional passives), there are no effects of heritability. When information in the environment is lacking (non-actional passives), then internal mechanisms kick in and heritability shoots up.

References

Babyonyshev, M., Ganger, J., Pesetsky, D., & Wexler, K. (2001). The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry, 32*(1), 1-44.

Borer, H., & Wexler, K. (1987). The maturation of syntax. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 23-172). Dordrecht: Reidel.

Borer, H., & Wexler, K. (1992). Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory, 10*, 147-189.

Brooks, P. J., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology, 35*(1), 29-44.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics, 15*, 467-473.

Demuth, K. (1989). Maturation and the acquisition of the Sesotho passive. *Language, 65*, 56-80.

Embick, D. (2004). On the structure of resultative particples in English. *Linguistic Inquiry, 35*(3), 355-392.

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition, 36*, 227-254.
MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of Child Language, 12*, 271-296.
Maratsos, M., Fox, D. E. C., Becker, J. A., & Chalkley, M. A. (1985). Semantic restrictions on children's passives. *Cognition*, *19*, 167-191.
Perlmutter, D. M., & Postal, P. M. (1984). The 1-advancement exclusiveness hypothesis. In D. M. Perlmutter & C. Rosen (Eds.), *Studies in relational grammar*.Chicago: University of Chicago Press.
Pinker, S., LeBeaux, D. S., & Frost, L. A. (1987). Productivity and constraints in the acquisition of the passive. *Cognition*, *26*, 195-267.
Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral genetics* (Fourth ed.). New York: Worth.
Strassberg, M., Peters, K., Marazita, M., Ganger, J. B., Watt-Morse, J., Murrelle, L., et al. (2002). Pittsburgh registry of infant multiplets. *Twin Research, 5*(499-501).
Wasow, T. (1977). Transformations and the lexicon. In P. W. Cullicover, T. Wasow & A. Akmajian (Eds.), *Formal syntax*.New York: Academic Press.
Wexler, K. (2005). Theory of phasal development. In *Plato's problem: Problems in language acquisition* (Vol. 48). Cambridge, MA: MIT Working Papers in Linguistics.