MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT

IV: RESULTS FOR STUDIES 2 AND 3: THE UKBTAT MODEL AND ITS APPLICATION TO NONTARGET ADDITIONAL LANGUAGE LEARNERS

Caroline Floccia, Thomas D. Sambrook, Claire Delle Luche, Rosa Kwok, Jeremy Goslin, Laurence White, Allegra Cattani, Emily Sullivan, Kirsten Abbot-Smith, Andrea Krott, Debbie Mills, Caroline Rowland, Judit Gervain, and Kim Plunkett

This article is part of the issue "Vocabulary of 2-Year-Olds Learning English and an Additional Language: Norms and Effects of Linguistic Distance" Floccia, Sambrook, Delle Luche, Kwok, Goslin, White, Cattani, Sullivan, Abbot-Smith, Krott, Mills, Rowland, Gervain, and Plunkett (Issue Authors). For a full listing of articles in this issue, see: http://onlinelibrary.wiley.com/doi/10.1111/mono.v83.1/issuetoc.

The aim of Study 2 was to develop the UKBTAT (UK Bilingual Toddlers Assessment Tool), the first screening tool for assessing the vocabulary size of bilingual 2-year-olds, in this case learning British English and 1 of the 13 target Additional Languages. The aim of Study 3 was to establish the reliability of UKBTAT for bilingual children learning British English and any other nontarget Additional Language. Norms for bilingual vocabulary in English were obtained through regression equations in linear mixed models, using variables shown to be predictive of comprehension and production in the previous analyses (Chapter III). Similarly, norms for the Additional Language vocabulary were calculated for children learning 1 of the 12 target Additional Languages (Spanish is sadly absent from the final list of target languages due to some disagreement with the Spanish CDI editors, TEA Ediciones).

Corresponding author: Caroline Floccia, School of Psychology, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK, email: caroline.floccia@plymouth.ac.uk
DOI: 10.1111/mono.12351

STUDY 2: CHARACTERISTICS OF THE UKBTAT

To be made freely accessible online at www.psy.plymouth.ac.uk/UKBTAT for professionals working with young children and academics, the UKBTAT is similar to the platform used for data collection in this project, but with modifications suited to an applied setting. Firstly, access is secured for practitioners or academics through a personal account, allowing confidential storage of patient test results. Practitioners can use the system to send a link to parents requesting the completion of tests, and have full access to all responses if required. Alternatively tests can be printed and used offline with parents. The tests are still presented in the same order as in this study, with English 100-word Oxford Short Form CDI, full Additional Language CDI (when Additional Language is one of the supported target languages), finishing with the Plymouth Language Exposure Questionnaire. The Plymouth Language Exposure Questionnaire is still the last component that must be filled in by the practitioner, either on the phone or in a live interview with the parent(s). Many of the questions from the family questionnaire and the Plymouth Language Exposure Questionnaire have been merged in an abbreviated Plymouth Language Exposure Questionnaire, which retains only the questions relevant to the significant predictors (amount of exposure, gender, overheard speech).

In the UKBTAT all children are assessed in English, with those whose Additional Language is one of the 12 target languages also assessed in their Additional Language. These data are used to calculate a percentile score for the child's position in their cohort for expressive and receptive vocabulary. For children whose Additional Language is assessed, separate ratings are provided for each language, otherwise only a single rating is provided for English.

STUDY 2: UKBTAT PREDICTIVE EQUATIONS

In order to be included in the UKBTAT model, predictors were required to reach significance in the ANCOVAs and the subsequent linear mixed models, and have an effect size of at least $\eta^2 = .02$ in the ANCOVAs (see Chapter III). With these criteria, predictors that made it through Step 1 were the relative amount of English child-directed speech (LEQ), the proportion of English in parental overheard speech (Overheard speech) and gender. No predictor relating to SES, the source of each language, the properties of the input and the status of the Additional Language survived Step 2 analyses. From the outset, we ruled out including Language Distance predictors in the equations since our aim was to provide models applicable to any Additional Language (Study 3).

Coefficients were obtained from the final mixed models shown in Table 18. In the case of English, these were obtained from a model run on the full cohort of 430 children, and on the 100-word Oxford Short Form CDI data rather than the 30-word CDI, in order to improve representativeness for the UKBTAT implementation (as mentioned before, the 30-word CDI shows a ceiling effect for a third of children in English comprehension). For the Additional Languages, the coefficients were obtained from models run on the 372 children who provided Additional Language data, and on the 30-word CDIs.

Altogether, these equations provide predicted scores for a bilingual of unspecified Additional Language in English, and in the Additional Language if it is part of our 12 target languages (Table 19).

For example, an Italian-English girl has 50% exposure to English (LEQ = 50) with parents speaking English and Italian equally often between themselves (Overheard speech = 3; this variable uses a 5-point scale, with 1 = parents always use the Additional Language when addressing one another, 2 = usually the Additional Language, 3 = English about half the time, 4 = usually English, 5 = always English; see Appendix 2). Using the

TABLE 18

Coefficient Estimates From the Linear Mixed Models for the UKBTAT, for Comprehension and Production, in Each Language (English and the Additional Language)

	Coef.	*SE*	*t*
English comprehension			
Intercept	37.62	3.78	9.97
LEQ	0.23	0.08	2.78
Overheard speech	5.19	0.90	5.76
Additional Language comprehension			
Intercept	21.83	1.87	11.70
LEQ	−0.06	0.02	−3.09
English production			
Intercept	34.48	5.76	5.99
LEQ	0.24	0.09	2.68
Overheard speech	3.85	1.14	3.39
Gender	−12.16	2.44	−4.98
Additional Language production			
Intercept	18.69	2.10	8.91
LEQ	−0.07	0.02	−3.71
Gender	−3.17	0.83	−3.80

Note. UKBTAT = UK Bilingual Toddlers Assessment Tool.
Dependent variables are scores on the 100-word CDI in English (comprehension: CDI100comp; production: CDI100prod) and on the 30-word CDI in the Additional Language (comprehension: ALCDI30Comp; production: ALCDI30Prod). For English, the models were calculated with $N = 430$ children and with $N = 372$ for the Additional Language. Variables were not *z*-scored so that they could be directly applied to new raw scores.

TABLE 19

Coefficients From the Linear Mixed Models of Vocabulary Knowledge in English and the Additional Language, in Comprehension and Production, and Decrements Derived From Standard Deviations of Residuals

	Coefficients				Decrements	
	Intercept	LEQ	Overheard Speech	Gender	10th Percentile Decrement	15th Percentile Decrement
Comprehension						
Predicted English 100 score	37.62	0.23	5.19		27.16	21.96
Predicted AL 30 score	21.83	−0.06			7.74	6.26
Production						
Predicted English 100 score	34.48	0.24	3.85	−12.16	28.55	23.09
Predicted AL 30 score	18.69	−0.07		−3.17	9.74	7.88

Note. For example, the predicted English score in comprehension is: $37.62 + 0.23 \times \text{LEQ} + 5.19 \times$ Overheard speech, with LEQ ranging from 0 to 100 (proportion of exposure to English vs. the Additional Language in child directed speech), and Overheard speech ranging between 1 and 5 (1 = parents always speak the AL between them; 2 = parents usually speak the Additional Language between them; 3 = parents speak the Additional Language and English half of the time; 4 = parents usually speak English between them; 5 = parents always speak English between them). Gender is assigned a value of 1 for girls and 2 for boys. Decrement is then applied to the predicted score, and compared to the observed score to determine if the child's score is below the 10th or the 15th percentile.

coefficients from the row label "Predicted English 100 score" in Table 19, the equation to be used would be $37.62 + 0.23 \times \text{LEQ} + 5.19 \times$ Overheard speech. Replacing LEQ with 50 and Overheard speech with 3, the child should have a predicted English score of 64.7 in comprehension, meaning she should understand 65 words from the 100-word Oxford Short Form CDI. Similarly, she obtains a score of 45.9 in English production using the appropriate equation in Table 19, meaning she would produce 46 out of the 100 words from the English Short Form CDI. In the Additional Language she obtains 18.83 in comprehension and 12.02 in production, meaning she should understand 19 words in Italian from the 30-word Italian CDI and produce 12.

As a diagnostic tool, it is important to be able to interpret the difference between a child's predicted and observed scores as a percentile. A reasonable threshold for suspecting a language delay is a score within the 10th percentile (Fenson et al., 2007; Rescorla, 2002; Tomblin, Records, & Zhang, 1996), so access to these ratings will allow practitioners to make an informed decision as

to whether a referral might be necessary in a near future, or whether a wait and see approach is more appropriate.

We therefore examined the distribution of observed–predicted residuals in the mixed models from which the coefficients above were generated. Standard deviations of these residuals were as follows: English comprehension 21.19, English production 22.28, Additional Language comprehension 6.04, Additional Language production 7.60. These provide a basis for converting an observed–predicted difference scored in items into a percentile, which is what UKBTAT reports for a child screened by this tool. Thus in English comprehension, an observed–predicted decrement of 21.96 items places a child at exactly the 15th percentile, and a decrement of 27.16 items at exactly the 10th percentile. In the example above, if the Italian-English girl who was predicted an English comprehension score of 64.7 words scored in reality less than 42.7 (i.e., 64.7–21.96), she would be in the 15th percentile (see Table 8).

Figure 4 provides the percentiles of word comprehension and production in English, illustrating the gender effect in production, and the well-documented difference between bilingual scores and monolingual data.

It is worth pointing out that a common practice recommended by Rescorla (1989), known as the Delay 3 cutoff, is to refer for further assessment any (American English) monolingual 2-year-old child who produces fewer than 50 words from the LDS (which contains 310 words, therefore 16%), which identifies about 15% of children. Our findings clearly point to the infeasibility of this "one size fits all" approach for bilingual children: the English vocabulary production score that would be needed to be in the 15th percentile or below varies between 0 and 42.3 (out of 100), depending on the extreme values of the predictors. That is, a boy with the minimum exposure to English, and whose parents would always speak the Additional Language

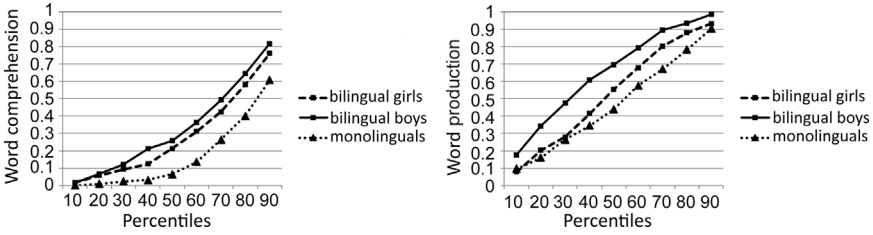

FIGURE 4.—Percentile values in English comprehension (left) and production (right) measured from the Oxford Short Form CDI, as function of gender, in the full cohort of bilingual toddlers ($N = 430$). For comparative purposes, monolingual data for comprehension and production are included on the same graphs ($N = 125$, taken from the Oxford CDI database, Hamilton et al., 2000). For example, in comprehension, children who understand a maximum of 50% of the CDI (*y*-axis) constitute about 75% of bilinguals, and 85% of monolinguals (*x*-axis).

between them, should produce on average 14.0 words out of 100 on the English CDI, and a score of 0 would put him on the 15th percentile. In contrast, a girl hearing 100% English as measured by the LEQ and whose parents always speak English between them, should produce 65.6 on the English CDI, and a score of 42.3 would place her on the 15th percentile.

STUDY 3: TESTING THE PREDICTIVE MODEL

We tested the validity of the equations above by seeing how accurately they could be used to predict novel data, namely those of the 58 nontarget Additional Language children, which was the aim of Study 3. Predicted scores for these children were generated using the equations (recalculated for the 372 children learning a target Additional Language) and then correlated with the observed scores. This was done for English scores only since nontarget children had supplied no Additional Language data. For both comprehension and production, strong correlations were seen between predicted and observed scores (comprehension: $r = .60$; production: $r = .59$). Importantly, there was not any systematic underprediction or overprediction of scores in these novel data, as established by t tests of the means of observed and predicted scores ($t < 1$).

We also calculated the number of children who would be identified as having a delayed language acquisition, by applying the conservative criterion of scoring less than 1 *SD* (16th percentile) below the mean of the overall distribution (Conti-Ramsden, Botting, & Faragher, 2001). For each child in the 58 nontarget Additional Language learners, using the standard deviations of residuals calculated above (UKBT predictive equations derived from the 372 children leaning the target Additional Languages), we converted the difference between each predicted score and the observed score as a percentile. Out of the 58 children, 9 were at or below the 16th percentile (15.5%) in comprehension and 11 in production (19.0%). Five out of the nine children with low comprehension scores had production scores at or below the 16th percentile, with the other four scoring also relatively low in production. Two children scoring very low on production had normal comprehension scores. Given the prevalence of 7–15% of experiencing delayed language acquisition (Kohnert, 2010) in monolingual children, these results suggest a very satisfactory sensitivity for the UBTAT equations.

This demonstrates that the model was not simply fitted to UKBT data a posteriori, but also has predictive validity, not only to fit a new set of data, but also to identify children with potential language delays. Furthermore, as the key test of the validity of our norms was carried out with data from children outside of the range of Additional Languages used to develop the model, it shows that the model is predictive of *general* bilingual vocabulary, whatever

Additional Language is spoken by the British-English learning bilingual. We therefore conclude that we have developed bilingual norms for 24-month-olds learning British English and any Additional Language, which have clear cut-offs for referral once the proportions of English in child-directed input and parental overheard speech have been determined, fulfilling the objectives of Studies 2 and 3.