Bilingual Children’s Acquisition of English Inflection: The Role of Language Dominance and Task Type

Johanne Paradis¹, Antoine Tremblay¹ and Martha Crago²
¹University of Alberta and ²Université de Montréal

1. Introduction

Children who are learning two languages experience more variability in their input than monolingual children. Bilingual children receive less input in each of their two languages than monolingual age-mates, and this input is seldom equally balanced between them. Bilingual children may have their exposure to one or both languages restricted to certain contexts, interlocutors, and registers/dialects. Finally, shifts in these quantitative and qualitative variations of exposure can occur over time due to changes in family structure, childcare arrangements, schooling, or place of residence. Assuming that variation in input conditions will affect language development, researchers have specifically asked the following question: How robust/sensitive is morphosyntactic acquisition in the face of dual language learning? This question has been addressed from both constructivist and generative perspectives, focusing on both rates and patterns of acquisition (Gathercole, 2007; Müller & Hulk, 2001; Nicoladis, Palmer & Marentette, 2007; Paradis, Nicoladis & Crago, 2007; Pérez-Leroux, Pirvulescu & Roberge, in press). This study examined bilingual-monolingual differences in the acquisition of English inflectional morphology with the aim of testing the constructivist model of bilingual development put forward in Gathercole (2007).

Recent research comparing the acquisition rates of young bilingual and monolingual children (Gathercole, 2007; Nicoladis et al., 2007; Paradis et al., 2007; Pérez-Leroux et al., in press; Thordardottir, Rothenberg, Rivard & Naves, 2006) has yielded the following set of findings: Bilinguals lag behind their monolingual age-peers in their morphosyntactic development in the preschool and early primary school years. However, bilingual-monolingual differences diminish or disappear when bilinguals are compared to monolinguals in their dominant language. In addition, bilingual-monolingual differences are greater for difficult target structures than for those that are transparent in form-function mapping, higher in type frequency, or generally-speaking, early-acquired. These results suggest that while bilingual-monolingual differences are in evidence, the impact of bilinguals’ variable input interacts with other factors such that protracted, across-the-board delay is not characteristic of their development.

To illustrate these points, let us look closely at the results of Paradis et al. (2007). They studied the production of the past tense and its French equivalent, the passé-composé, in four-year-old French-English bilingual children. They found that the bilingual children were as accurate as their monolingual age-
mates with the regular past tense or passé-composé in their dominant language; however, English-dominant bilinguals were less accurate with the more difficult irregular past tense forms than English monolinguals. The “irregular” passé-composé forms in French are less difficult than those in English, and accordingly, French-dominant bilinguals performed akin to monolinguals for this structure. Furthermore, the lag between bilinguals and monolinguals was not entirely commensurate with differences in input frequency, i.e., these bilingual four-year-olds might have had half the exposure to English as monolinguals, but they did not exhibit a two-year delay in their accuracy levels with the past tense. Finally, the French-dominant bilinguals often performed in the range of monolinguals with specific language impairment (SLI) in English. This indicates that over-identification of language disorder in bilingual children this age is a distinct possibility, and so, bilingual-monolingual differences in acquisition rates have consequences for applied as well as theoretical domains. It is possible that the influence of dominance and target structure difficulty could diminish in French-English bilinguals by the time they are six years old. For instance, bilingual–monolingual differences might be less pronounced or disappear in the non-dominant language, for difficult target structures, or both. As a consequence, French-dominant bilinguals might cease to overlap with monolinguals with SLI in their abilities with English inflection. One objective of this study was to examine this possibility.

Prior research comparing bilingual and monolingual rates of morphosyntactic development has been based on either production data or grammaticality judgment data alone (e.g., Gathercole, 2002; Nicoladis et al., 2007; Paradis et al., 2007), or production and comprehension data compared to monolingual norms (Thordardottir et al., 2006). Gathercole & Thomas (2005) examined both production and comprehension data, but among bilinguals only. To the best of our knowledge, no studies to date have examined bilingual and monolingual children’s abilities to produce correctly, and to judge correct use of, the same target structure at the same age. Understanding whether bilingual-monolingual differences would change as a result of task type is important for evaluating the adequacy of theoretical claims (see below). Furthermore, grammaticality judgment tasks are not equivalent to comprehension tasks, and therefore, provide additional information about language development.

It is commonly understood that comprehension-style receptive tasks are less demanding than expressive tasks, and thus, children often show superior knowledge on the former than the latter, for the same structure. For example, toddlers display knowledge of verb argument structure through comprehension tasks before they regularly use these structures in their spontaneous speech (Hirsh-Pasek & Golinkoff, 1996). Grammaticality judgments, on the other hand, demand a high level of metalinguistic awareness because they demand reflection on (implicit) grammatical knowledge rather than meaning. Thus, even though grammaticality judgments typically require a minimal response like a comprehension task, they are arguably more demanding than comprehension tasks. Furthermore, dual language learning may confer superior executive
control functions on young bilingual children compared with their monolingual age-mates (Bialystok, 2007). Certain metalinguistic abilities can be conceptualized as part of executive control. This cognitive difference between bilinguals and monolinguals could cause bilinguals’ performance, vis à vis monolinguals, to be different when given a grammaticality judgment task versus a production task, for the same target structure.

1.1 Gathercole’s (2007) constructivist model of bilingual acquisition

Constructivist (or usage-based) accounts of the lexicon and morphological acquisition assume that lexical items are stored multi-morphemically, and that what are commonly construed as “rules” for inflectional affixation, like [verb+ed] = verb(past tense), are actually schemas that emerge across the numerous stored types of [verb+ed] concatenations in the lexicon. Schemas of this kind are item-specific, and hence, language-specific. A constructivist approach assumes that abstract levels of grammatical representation, such as the general transitive construction template - subject-verb-object, emerge from numerous types of more specific constructions, such as transitive phrases with lexical material present, stored in the lexicon. But, a constructivist view of linguistic representation, even for the endstate grammar, is much less abstract than most generative views (Tomasello, 2003). Regarding inflectional morphology, it is likely that multi-morphemic lexical forms and their schemas remain specific, meaning there is a conflation of morphological form and abstract knowledge of that form (Bybee, 2001; Gathercole, 2007). In other words, the most abstract level attained might be the level at which speakers can perform a “wug” test and add a certain inflection to a novel base in production. And the ability to produce a correct morphological form and to judge its well-formed production by others would draw on the same underlying source. This is a point of contrast with most generative approaches, where abstract principles delineating well-formedness in morphosyntax can be disassociated from the morphophonological details of individual items in the lexicon (see, for example, Wexler, 1998).

On a constructivist account, acquisition of morphological schemas is considered to be piecemeal, meaning tied to individual lexical forms at first, and to be sensitive to input frequency to reach full productivity. Morphological forms with high type frequency and transparent form-to-function mapping require less exposure to become productive than more idiosyncratic and complex forms. This approach is compatible with the findings that bilinguals lag behind monolinguals in their accuracy with morphology until they have achieved a critical mass of input to neutralize differences, and furthermore, differences are more protracted for more complex forms (Gathercole, 2007). The assumption that morphological schemas are item- and language-specific leads to two additional predictions: That bilinguals will acquire their two languages rather autonomously in this domain, such that, inflectional morphology in the language of greatest exposure could be acquired faster, and that interaction in
the form of transfer would be nonexistent at the level of inflectional morphology (Gathercole, 2007). This last assumption rules out “carry-over” or bilingual bootstrapping in this domain of acquisition. In contrast, Gathercole’s (2007) model predicts interaction for domains of language that implicate shared cognition, i.e., conceptual and pragmatic systems, although, mainly in balanced bilinguals and late in the developmental process.

1.2 Research questions for this study

This study was designed to address the following questions; (1) Would dominance and target structure difficulty determine bilingual-monolingual differences in six-year-olds? Would bilingual six-year-olds’ abilities be confounded with those of a clinical population? Addressing this question was intended as a follow-up to Paradis et al., (2007)’s study with four-year-old children from the same population. (2) Would bilingual-monolingual differences, if any, change depending on whether a production or grammatical knowledge task were given to the children? On the one hand, a strict interpretation of Gathercole’s (2007) model leads to the prediction that bilingual-monolingual differences would be similar for these two tasks because the children’s productive and receptive abilities would come from the same, item- and language-specific sources. On the other hand, consideration of the superior executive control abilities of bilingual children, and the exigencies of bilingual speech production, could predict task-based differences in their performance with the same target structure.

2. Method

Forty-four French-English bilingual children in grade one in French mother tongue schools in Edmonton, Canada (an English majority city), participated in this study. Children’s dominant language was assessed through standardized vocabulary tests on the grounds that greater exposure to a language would result in a larger vocabulary in that language. Children were assigned to the English-dominant or French-dominant group depending on which of their standard scores on the PPVT (Dunn & Dunn, 1997) and the EVIP (French equivalent of PPVT; Dunn, Thériault-Whalen & Dunn, 1993) was higher. This resulted in 24 French-dominant (mean age = 6;10) and 20 English-dominant (mean age = 6;9) children. A t-test confirmed no difference in the ages. This dominance classification based on vocabulary is significantly correlated with our classification of children’s home language exposure (mainly English, mainly French, or both French and English) based on parental questionnaires ($r = .611$, $p = .01$). Monolingual English-speaking children’s performance on the dependent measures was obtained from the norming sample means from the Test of Early Grammatical Impairment Examiner’s Manual (TEGI: Rice & Wexler, 2001).
The TEGI consists of elicitation and grammaticality judgment probes. For this study, children were given the third person singular (THIRD SING) and past tense (PAST TNS) elicitation probes. For the THIRD SING probe, children were shown a series of pictures of professionals engaged in typical activities and given prompts like “here is a teacher. Tell me what a teacher does”, with an expected response of “a teacher teaches” or “a teacher writes on the board”. For the PAST TNS probe, the children were shown a series of picture pairs, where in the first picture a child is engaged in an activity, and in the second picture the activity has been completed. Children were given prompts like “here the boy is raking. Now he is done. Tell me what he did?”, with an expected response of “he raked”. For the grammaticality judgment probes, the experimenter acted out a scenario with two toy robots who were learning English, and asked the children to judge if their sentences were “said right” or “not so good”. Sentences included three types of grammatical errors, which constitute three probes after scoring: (1) DROP TNS (he running away, he look happy), (2) BAD AGR (he am hurt, I drinks milk), (3) DROP ING (he is smile, the bear is jump). The target structures for errors and well-formed sentences within each probe include both bound and unbound (Be) morphology. This contrasts with the production probes we used that only target bound morphology. We take this into account in our analyses. It is important to keep in mind that because we are using the TEGI norming sample as a monolingual control group, it was necessary to administer the grammaticality judgment probes exactly as instructed in the manual, which meant including targets for Be.

Note that the BAD AGR and DROP ING errors are errors with grammatical morphology, but they are not errors of omission of tense-marking morphology like DROP TNS errors. Both the THIRD SING and PAST TNS probes target the production of tense-marking morphology. Much research has shown that tense-marking morphology is late-acquired in English (see Rice & Wexler, 2001, for review). Explanations for why finite verb morphology would be a difficult structure in English acquisition have been put forward from both a generative (Wexler, 1998) and a constructivist perspective (Theakston, Lieven & Tomasello, 2003; Bybee, 2001), although Wexler’s (1998) proposal specifies tense-marking morphemes, and omission errors, in particular. For this study, the THIRD SING, PAST TNS and DROP TNS probes are construed as tests of difficult-to-acquire structures, while the BAD AGR and DROP ING probes are not.

3. Results

The mean proportion correct scores on the THIRD SING and PAST TNS probes for the French-dominant and English-dominant bilinguals are presented in Figure 1 along with the monolingual norming sample means for children with typical language development (TLD), and those with SLI, for the same age range. A two-way mixed ANOVA with dominance (French, English) as an independent groups factor, and morpheme type as a repeated measure was
performed on the scores of the bilinguals. There was a significant main effect for dominance group ($F(1,38) = 12.26, p = .001$, partial eta squared = .244), but no other effects were significant. Thus, the English-dominant bilinguals were more accurate than the French-dominant bilinguals for both morpheme types equally. One-sample t-tests comparing the bilinguals to the monolingual groups revealed the same pattern for THIRD SING and PAST TNS. English-dominant bilinguals were as accurate as the monolinguals with TLD, but more accurate than monolinguals with SLI for both morpheme types (THIRD SING: $t(19) = 15.27, p = .000$; PAST TNS: $t(18) = 11.47, p = .000$). The French-dominant bilinguals were less accurate than the monolinguals with TLD (THIRD SING: $t(22) = -4.19, p = .000$; PAST TNS: $t(20) = -4.29, p = .000$), and were as accurate as the monolinguals with SLI for both morphemes.

In Figure 2, the bilingual and monolingual children’s performance on the grammaticality judgment probes is shown, divided according to error type. Performance was measured using A-prime scores. A-prime scores are designed to correct for the “yes” bias that young children often have in these tasks. Group means for the monolinguals on the TEGI are given as A-prime scores. A two-way mixed ANOVA with dominance group (French, English) as an independent groups factor and error type as a repeated measure (DROP TNS, BAD AGR, DROP ING) was performed on the bilinguals’ scores. There was a significant main effect for error type ($F(2,82) = 10.14, p = .000$), for dominance ($F(1,41) = 5.73, p = .021$, partial eta-squared = .123), but no significant interaction. Therefore, the English-dominant children had higher scores regardless of error type than the French-dominant children. LSD pairwise comparisons of the estimated marginal means revealed that DROP TNS (.85) was significantly
lower than either BAD AGR (.92) or DROP ING (.94), but the latter two were equivalent. Turning to comparisons with the monolingual groups, one-sample t-tests revealed that the English dominant bilinguals performed identically to the monolinguals with TLD for DROP TNS and DROP ING, but had higher scores for BAD AGR ($t(19) = 2.95$, $p = .008$). The English dominant bilinguals also had higher A-prime scores than the monolinguals with SLI for each probe (DROP TNS: $t(19) = 12.48$, $p = .000$; BAD AGR: $t(19) = 35.4$, $p = .000$; DROP ING: $t(19) = 21.1$, $p = .000$). The French dominant bilinguals had lower A-prime scores than the monolinguals with TLD for DROP TNS ($t(22) = -3.21$, $p = .004$) and BAD AGR ($t(22) = -2.43$, $p = .024$), but not for DROP ING. In contrast to their results for production, the French dominant bilinguals had higher A-prime scores than the monolinguals with SLI (DROP TNS: $t(22) = 4.59$, $p = .000$; BAD AGR: $t(22) = 3.39$, $p = .003$; DROP ING: $t(22) = 5.33$, $p = .000$).

Because the DROP TNS and BAD AGR error types included targets with lexical verbs, as well as targets with Be, we wanted to determine whether these patterns held when only the lexical targets were considered. This is important for comparisons with the production probe results. Children’s lexical A-prime scores were slightly lower than the lexical and Be combined scores for DROP TNS (English dominant: .87 vs. .91; French dominant: .75 vs. .80) and BAD AGR (English dominant: .98 vs. .99; French dominant: .82 vs. .85). However, the key patterns remained the same when the above analyses were re-run with the lexical-target-only scores: Children had significantly higher scores for BAD AGR than DROP TNS, and the French dominant bilinguals had higher A-prime scores than the monolinguals with SLI on these probes.
We have not yet compared children directly on the production and grammaticality judgment probes, but instead have compared them indirectly through interpreting their performance with respect to monolinguals. This is because proportion correct in production is not measured using the same scale as A-prime. We pursued a more direct approach by calculating the proportion of the bilingual children in the study who performed at or above the criterion score appropriate for their age on the TEGI probes. The results of this calculation are in Figure 3. This analysis reveals that more bilingual children met monolingual age-expectations for the grammaticality judgment probes (.65, .80, .75) than the production probes (.56, .53). Looking at the results according to dominance groups, the difference between the proportion of French vs. English dominant children who met age expectations was much smaller on the grammaticality judgment probe DROP TNS (.15) than on the production probes THIRD SING (.24) and PAST TNS (.39).

For an additional comparison between the grammaticality judgment and production tasks, we examined differences between the means of the dominance groups. Differences in the mean proportion correct and A-prime scores between English and French dominant bilinguals were less than half for DROP TNS than the production probes (.18 vs. .28, .28), and Cohen’s d effect sizes were moderate for the differences in the production probes (.438 and .497), but small for the grammaticality judgment probe (.303). Parallel comparisons were carried out between the children’s mean scores on the two production probes and the grammaticality judgment probe DROP TNS, for the lexical-target-only scores, and produced similar patterns.
Finally, we conducted a series of Pearson correlations between the bilingual children’s scores on the two production probes and grammaticality judgment probe DROP TNS and DROP TNS (lexical only). Correlation coefficients ranged from .714 to .798 between probe types, all significant at \( p < .01 \). These moderate-to-large correlations suggest that children were drawing on similar sources of knowledge to perform on these tasks.

4. Discussion

We set out to investigate whether dominance, target structure difficulty and task type would influence bilingual children’s accuracy with English inflectional morphology, vis à vis their monolingual peers. Our results showed that at an age when monolinguals with TLD have reached ceiling in their performance on the measures used, English-dominant bilingual children can do so as well, and even when English is not their language of schooling. The interaction of dominance and target structure difficulty was shown in particular by the narrowed gap in performance between English- and French-dominant bilinguals on BAD AGR and DROP ING compared with DROP TNS. These results are consistent with findings from younger bilinguals in the same social context (Paradis et al., 2007) and bilinguals the same age in other contexts (Gathercole, 2007). In addition, comparisons between the production probes and the DROP TNS grammaticality judgment probe revealed a third interactive factor: task type. In a nutshell, this study found that bilingual-monolingual differences were (1) more concentrated in the production of difficult-to-acquire morphemes than in the judgments of well-formed use of these morphemes, and (2) non-existent at this age in bilinguals’ dominant language. These findings reinforce the notion that even though bilingual acquisition is sensitive to these children’s variable input, this sensitivity is mitigated by intervening factors. While two of these factors, dominance and target structure difficulty, could be derived from input variability (Gathercole, 2007), it is not obvious that task effects could be.

In Section 1.2, we predicted that Gathercole’s (2007) model would be most compatible with equal performance in production and grammaticality judgment tasks. This is because abstract morphological representation and morphophonological details in the lexicon would be largely conflated. That children were drawing on a common knowledge base to perform on these measures was suggested by the correlational analysis. And yet, these bilingual children performed closer to their monolingual peers in their non-dominant language on the grammaticality judgment probes than the production probes. One explanation for this discrepancy could be that abstract knowledge of obligatory finiteness marking could be carried over from the dominant language. We do not favour this explanation for two reasons. First, as stated above, on Gathercole’s (2007) model, the two languages of a bilingual would not be expected to interact at the level of item-specific, language-specific schemas. Second, prior research on bilingual acquisition has found no evidence for
crosslinguistic influence at the level of finite verbal morphology (Hulk & Müller, 2000; Paradis & Genesee, 1996).

A more promising explanation lies in the contrastive metalinguistic, or more precisely, executive control abilities of bilinguals and monolinguals. The performance gap between bilinguals and monolinguals could be narrowed on a grammaticality judgment task because bilingual children might be more efficient at reflecting on their knowledge of correct forms than at producing the same correct forms. Production entails accuracy in the execution of fine-grained morphophonological details, but well-formedness judgments do not. Automaticity in accessing and pronouncing stored forms in real-time speech production would be sensitive to the amount of practice with a language, and bilinguals would have less than monolinguals. Furthermore, competition between the two languages of a bilingual is possibly fiercer in production than it is for other linguistic tasks, and thus, production requires more attentional resources, in the non-dominant language in particular (cf. Bialystok, 2007). Conversely, on certain grammaticality judgments tasks, bilinguals have been shown to have superior performance to monolinguals, which has been attributed to their superior executive control functions (Bialystok, 2001, 2007).

We are not arguing that bilinguals are generally superior to monolinguals in grammaticality judgment tasks, as indeed our data are not consistent with that claim, and prior research has shown conflicting results supporting this claim (Bialystok, 2001, 2007). If bilinguals perform worse than monolinguals on a grammaticality judgment task, this most likely means that they have accumulated less knowledge of the target structure (cf. Gathercole, 2002). Instead, we propose that production tasks might underestimate the knowledge bilinguals possess (as referenced to monolinguals) because they are additionally demanding for bilingual children, and grammaticality judgments might provide a more accurate picture of their accumulated knowledge (again as referenced to monolinguals) because bilingual children have a facility for them. To turn the argument around, in monolingual children, differences between production and grammaticality judgment tasks as measures of the same linguistic knowledge would be expected to be smaller than in bilinguals.

Finally, the results of this study have applied implications for educators and speech language pathologists. Paradis et al. (2007) found that over-identification of SLI might occur in bilingual four-year-olds and the present study indicates that it could also occur in bilingual six-year-olds. But, the potential for misidentification would vary according to whether a child’s dominant language was being assessed, the difficulty of the target structure being probed, and the type of task used to determine linguistic knowledge. In other words, it is not the case that bilinguals are merely, say, two years behind monolinguals across-the-board, but instead they have a unique profile. The presence of this unique profile underscores the need for bilingual norm referencing for standardized tests (see also Thordardottir et al., 2006).
References


