Cross-linguistic comparisons in child language research

RUTH A. BERM AN
Tel Aviv University

ABSTRACT

Major large-scale research projects in the early years of developmental psycholinguistics were English-based, yet even then numerous studies were available or under way in a range of different languages (Ferguson & Slobin, 1973). Since then, the field of cross-linguistic child language research has burgeoned in several directions. First, rich information is now available on the acquisition of dozens of languages from around the world in numerous language families, spearheaded by the five-volume series edited by Slobin (1985–1997) and complemented by in-depth examination of specific constructions—e.g. causative alternation, motion verbs, passive voice, subject elision, noun compounding—in various languages, culminating in an in-depth examination of the acquisition of ergativity in over a dozen languages (Bavin & Stoll, 2013). A second fruitful direction is the application of carefully comparable designs targeting a range of issues among children acquiring different languages, including: production of early lexico-grammatical constructions (Slobin, 1982), sentence processing comprehension (MacWhinney & Bates, 1989), expression of spatial relations (Bowerman, 2011), discourse construction of oral narratives based on short picture series (Hickmann, 2003) and longer storybooks (Berman & Slobin, 1994), and extended texts in different genres (Berman, 2008). Taken together, research motivated by the question of what is particular and what universal in child language highlights the marked, and early, impact of ambient language typology on processes of language acquisition. The challenge remains to operationalize such insights by means of psychologically sound and linguistically well-motivated measures for evaluating the interplay between the variables of developmental level, linguistic domain, and ambient language typology.

A personal note to start: raised and educated in English, immigration to a Hebrew-speaking country at the outset of my career led me into the domain of ‘contrastive analysis’ in foreign language pedagogy (e.g. Fries, 1945;
CROSS-LINGUISTIC COMPARISONS IN CHILD LANGUAGE RESEARCH

Lado, 1957; Wardaugh, 1970). My work on ‘English for speakers of Hebrew’ in the 1960s paved the way for concern with Modern Hebrew structure in the 1970s, subsequently with acquisition of Hebrew as a first language, in which the Journal of Child Language hosted my initial publications (Berman 1981a, 1981b, 1982). In this homage to the journal, and to the field, I focus on contrastive analysis from the perspective of typologically motivated cross-linguistic research on child language over the past four decades.

The early years of developmental psycholinguistics in the 1960s and 1970s was dominated by a largely English-language orientation, reflected in three complementary projects devoted to the development of syntax at Harvard, Berkeley, and Maryland (Ferguson & Slobin, 1973, pp. 295-521), culminating in Roger Brown’s pivotal (1973) book *A First Language: The Early Stages*. Yet from the start, the field was not strictly anglocentric. Ferguson and Slobin’s anthology includes studies from some dozen languages other than English, with Slobin’s groundbreaking chapter on cognitive prerequisites listing “available material on the acquisition of 40 different native languages” with some dozen additional languages “currently in progress” (1973, pp. 177-79). Notably, the first volume of *JCL* in 1974 includes studies on several different languages, for example, Finnish (Bowerman, reviewed by Griffith), German (Tschang-Zin Park), Serbocroatian (Savic & Mikeš), and Tamil (Garman). The move beyond English received a tremendous boost in the 1980s and 1990s with the publication of the five volumes edited by Dan Slobin on *The Crosslinguistic Study of Language Acquisition*—which included descriptions of first language acquisition of nearly thirty languages from over a dozen families, spanning all five continents. Cross-language comparability was facilitated by the fact that the bulk of the chapters in these volumes addressed a prescribed set of issues and followed a similar format.

One marked advance in the field, then, has been the accumulation of a vast store of knowledge on acquisition of both familiar and more exotic languages—including rich data based on ethnographic fieldwork in languages from diverse families (for example, Allen, 1996; Bavin, 1990; Brown, 2012; Demuth, 1984; Pfeiler, 2007; Pye, 1986; Suzman, 1991, 1999). A second significant development in the domain was the shift beyond description to the principled issue of the impact of first language typology on the process of acquisition, or the “overarching question ... [of] the extent to which the development of linguistic ... representations is influenced by the particular language the child is learning” (Slobin, Bowerman, Brown, Eisenbeiss & Narasimhan, 2011, p. 134).

Responses to this question have been expedited by two major developments in the field: (i) the establishment and continued extension
and updating of the Child Language Data Exchange System (CHILDES, MacWhinney, 2000, 2014; MacWhinney & Snow, 1990) which, as its name implies, gives researchers access to a rich digital database of children's speech in dozens of different languages, enabling them to freely compare across languages, children, and topics of interest; and (ii) the design of closely comparable methods of data elicitation among children at similar developmental stages in different languages.

The latter approach was first articulated by the 'field manual' edited by Slobin and his colleagues at Berkeley (Slobin, 1967), and subsequently applied in the framework of the Berkeley cross-linguistic acquisition project carried out in 1972–1973 in four typologically different languages (English, Italian, Serbo-Croatian, and Turkish) on the basis of a meticulous cross-sectional methodology tapping lexico-grammatical productions of children in eight age-groups between ages 2;0 and 4:4 (Ammon & Slobin, 1979; Johnston & Slobin, 1979; Slobin & Bever, 1982). This complex project yielded important insights into universal compared with language-particular features of acquisition of a range of morphosyntactic constructions, highlighting the interplay between typologically specific linguistic properties and shared cognitive developments in the process (Slobin, 1982).

A second major undertaking spearheaded by Slobin in the 1980s was the so-called ‘frogstory project’, in which children aged three to nine years compared with adults, native speakers of five different languages (English, German, Hebrew, Spanish, and Turkish), were asked to recount the contents of a picture-book story without words (Berman & Slobin, 1994). This project was methodologically significant, since it applied closely parallel procedures of elicitation, transcription, and analysis to cross-linguistic, developmental comparisons in four major domains of linguistic expression—tense-aspect, event representation, perspective-taking, and connectivity—and it addressed these issues in the context of extended (narrative) discourse, rather than on the basis of isolated words or sentences, on the one hand, or of child–adult conversational interaction, on the other.

As such, the frogstory research supplemented earlier work of Annette Karmiloff-Smith (1979) comparing acquisition of articles and pronouns in English and French and of Maya Hickmann (1980) and her associates analyzing reference to person, space, and time in four different languages (e.g. Hickmann, Hendriks, Roland & Liang, 1996), as summarized in Hickmann (2003). The Karmiloff-Smith and Hickmann studies elicited speech output on the basis of short, specially designed picture-series as compared with the longer, more elaborate series of events depicted in the Mercer Mayer’s Frog, Where are You? booklet. Besides, these scholars were initially concerned primarily with nominal reference rather than with the role of predicates in event representation at the core of our study.
Motivating the introduction of text-embedded narrative contexts adopted by these different projects was 'a functional approach to child language'. A form–function perspective also underlies cross-linguistic research on sentence processing conducted by Brian MacWhinney and the late Liz Bates (Bates & MacWhinney, 1982; MacWhinney & Bates, 1979) and their associates (e.g. Bates, McNew, MacWhinney, Devescovi & Smith, 1982), culminating in the Bates and MacWhinney (1989) collection. For example, their comparative investigation of Serbo-Croatian, Navajo, and American Sign Language led them to suggest "that, although the basic communicative function [of topicalization and commenting] is presumably universal, there is an enormous variation across languages in the particular surface devices or forms used to encode these functions". Their analyses of sentence processing in diverse languages, including in acquisition, by means of experimentally designed comprehension studies, played a key role in development of the 'competition model' of language learning, showing, for example, how cues representing typological features of the target language (e.g. animacy compared with word order) play differential roles in sentence comprehension by speaker–learners of particular languages.

In contrast to these largely comprehension-directed, sentence-level analyses, studies on children's discourse-embedded productions in different languages, like those based on picture-series or picture-book narratives, demonstrated the value of tracing the development of linguistic forms in contexts "beyond the sentence" (Berman, 2009a). In such research, a form–function approach (Slobin, 2001) addresses the question of how lexical and morphosyntactic constructions are deployed across development in different languages in relation to the functions they serve in extended discourse—in domains such as connectivity, event representation, perspective-taking, reference, or temporality. As shown by the Bates and MacWhinney work on sentence processing, these discursive functions, too, are largely shared across languages, but they are realized by typically language-particular forms that from early on have an important impact on developing child language. For example, the Berman and Slobin (1994) study showed that, developmentally, three-year-olds addressed the task of 'relating events in narrative' in a very different fashion than five-year-olds, and that these in turn differed markedly from the narratives produced by the nine-year-olds in different languages: across the languages, the youngest children tended to relate (to) the contents of each picture as isolated events, by late preschool age they were able to chain events and to mark their chronological sequence, while the schoolchildren had internalized a narrative schema in the shape of a shared action-structure based on causally related episodes, and adolescents and adults typically elaborated on plot-advancing events by means of rich evaluative devices and personal interpretations of the episodes they recounted. In cross-linguistic perspective, on the other
hand, from the youngest age on, participants produced texts that were constructed in keeping with the typologically unique features of each of the target languages, so that the task of becoming native speakers interacted importantly across age-groups with the task of becoming proficient narrators.

This ‘frogstory’ project spawned numerous parallel studies in a larger range of languages and in varied cultural settings (Strömqvist & Verhoeven, 2004), while also engendering a rich array of cross-linguistic comparisons in such domains as narrative development (so further complementing the work of Hickmann, 2003, and Karmiloff-Smith, 1979), on the one hand, and of motion verbs, on the other (Slobin, 1996a, 1996b, 2004a, 2004b), themes that are represented by the studies in Guo, Lieven, Ervin-Tripp, Budwig, Özçalışkan, and Nakamura, 2009, pp. 121–236). Slobin’s psycholinguistic extensions of Talmy’s (1985) typology of different types of figure-ground relations to child language acquisition showed that by early preschool age, children would encode motion events along the lines of the preferred means of expression and the linguistic repertoire of grammatically marked distinctions characteristic of their native language. For example, even the three-year-olds differed in how they described events of running away or falling from a tree or into a body of water—in verb-internal terms in Hebrew and Spanish compared with by means of locative particles in English or German. On the other hand, however, the English-acquiring children were more like their Spanish-speaking peers in how they distinguished aspectually ongoing from telic events, as compared with their German- or Hebrew-speaking counterparts.

The burgeoning analysis of children’s encoding of narrative events in different languages provided discourse-embedded evidence for Bowerman’s (1985, 1993) claim that, from the outset, children’s semantic categorizations are affected by linguistically specific form-meaning mappings in the ambient language. Rich data supporting children’s early sensitivity to language-specific encodings of different conceptual domains are provided by the cross-linguistic work of Bowerman and her associates on spatial reference, based on closely comparable experimental designs—typically requiring toddlers acquiring and adults speaking different native languages to describe the contents of pictures depicting specific event types such as cutting, putting, and taking. Such investigations of children’s encoding of spatial categories in different languages cover an ever-growing range of languages—some closely related like English and Dutch, others typologically far removed, like English and Korean (e.g. Bowerman & Choi, 2001, 2003; Choi & Bowerman, 1991; Choi, McDonough, Bowerman & Mandler, 1999; Majid, Bowerman, Van Staden & Boster, 2007; Majid, Gullberg, Van Staden & Bowerman, M., 2007; Narasimhan, Kopecka, Bowerman, Gullberg & Majid, 2012). Taken together, this body of research provides
that were robust, empirically anchored evidence for important insights into the relationship between language and cognition, such as Slobin's (1996a) conception of verbalization as a process of "thinking for speaking", and Bowerman's (2011) insistence that sensitivity to the typological specificities of the ambient language plays a key role in children's (re-)construction of their native language from early on, in the very first year(s) of life. And it effectively buried the idea (deriving largely from personal experience of second language learners) that a given language is somehow easier or harder for children to acquire—while raising the more interesting query as to which form-function mappings require more or less time and effort from children acquiring different native languages.

Another line of cross-linguistic research that evolved largely from the 'frogstory' studies shifted concern from early acquisition to later, school-age language development and use, a domain formerly considered primarily in clinical or pedagogical terms, rather than from a developmental psycholinguist perspective. One such project applied closely parallel procedures for eliciting extended texts—both narrative and expository, in both speech and writing—in seven different languages (Berman, 2008). Results of this richly complex design are written up in two special journal issues, each with a specific research goal: in the first, distributional findings are compared across the different languages in a range of domains—lexicon, noun phrase constructions, tense-aspect, passive voice, modality, clause-combining, and text openings and closings (Berman & Verhoeven, 2002); in the second, texts in each of the participating languages—Californian English, Lyonnais French, Iberian Spanish, Israeli Hebrew, Netherlands Dutch, Scandinavian Icelandic and Swedish—were analyzed around the shared topic of expression of discourse stance as relatively subjectively involved compared with more impersonally detached (Berman, 2003). Findings of this project reiterate the insight of both the experimental work of Bates and MacWhinney and the oral narrative studies of Slobin and his associates: on the one hand, they demonstrate shared discourse-based communicative functions across languages and children (e.g. early sensitivity to genre-appropriate use of linguistic forms; later development of global level principles of text construction in expository than narrative discourse; and high-school adolescence as a watershed in developing discourse abilities in different communicative settings); on the other hand, the particular linguistic forms and constructions serving to encode these functions differ markedly as a function of native language typology and culture (e.g. favoring of passive voice in texts written in English and Dutch compared with those in Hebrew and Spanish, as two languages that can rely on subjectless impersonal constructions; the more marked contrast between written and spoken registers of usage in texts produced in French and Hebrew compared with those in Swedish and English; and the differences in expression of a
RUTH A. BERMANN

generalized, impersonal stance in languages that have a specially dedicated impersonal subject pronoun like French and Swedish as against languages that readily host subjectless impersonal constructions like Hebrew and Spanish (Berman, 2011). A major, hitherto largely undocumented, insight from this project thus extends a key finding from the Berman and Slobin 'frogstory' analysis of oral, picturebook narratives among children aged three to nine years old—that linguistic forms “have a long developmental history”. Development of linguistic knowledge and language use in different communicative contexts is shown to extend well beyond nine-year-old middle children and even twelve- to thirteen-year-old pre-adolescence. In each of the languages in this project, flexibly proficient use of an extensive repertoire of linguistic forms consolidated only by high-school adolescence; and this occurred in tandem with the emergence of language-specific rhetorical preferences and culturally determined as well as individual styles of expression. A relevant conclusion, then, is that linguistic typology interacts with social-cognitive maturation and cultural norms in the path from early established 'native speakers' to the later emergence of proficiently literate language users.

What lessons and insights can be derived from forty years of cross-linguistic research on child language surveyed above? First, these studies have added both depth and breadth to the traditionally structuralist field of linguistic typology as well as to current more functionally motivated typological research by applying USAGE-BASED APPROACHES and psycholinguistic principles to the analysis of a diverse range of ‘child languages’. These studies also typically include comparisons with naturalistic adult usage (in the case of child–adult interactions) or take adult responses as representing a standard of proficient language use (in the case of semi-structured elicitations and experimental designs). As such, they aid in demonstrating not only what properties characterize a given language, but which of these exert a clear impact on language use from very early on and so play a key role in acquisition as against ones that are shared across children and across languages (Berman, 1986; Bowerman, 2011; Slobin, 1982).

Carefully designed cross-linguistic child-language studies have also extended and deepened our knowledge of SPECIFIC DOMAINS of linguistic structure and language use. These include issues of sentence processing, spatial reference, and motion events in work referred to earlier by scholars such as Bates and MacWhinney and by Bowerman and Slobin and their associates respectively. Such work is supplemented by in-depth analyses of acquisition of selected linguistic systems in different languages, for example: (a) causative constructions and transitivity alternations (Allen, 1998; Berman, 1993a, 1993b; Bowerman, 1974; Hochberg, 1986; Pye, 1993); (b) passive voice constructions (e.g. Allen & Crago, 1996; Demuth, Jisa, Reilly, 2002; Jakubowicz, 1999; Gelman, 1999). Direction to acquisition of familiar and also comparable child language use is often also evaluated by using naturalistic adult usage in the case of child–adult interactions or semi-structured elicitations and experimental designs. As such, it would appear that child language use is significantly shaped by the social-cognitive context of its use, and this is something that is true of adult language use as well.

As a cross-language comparison study, the Berkeley data base contains information on a large number of languages from around the world. (Evans & Demuth, 2006). These include a wide range of languages from different world language families, as well as a number of non-literate languages, and cover the major world language families. This information is available online at the Cross-Linguistic Data Repository (CLDR) and the Berkeley Language Data Base (BLDB). The data base is currently undergoing a major revision to make it more accessible to researchers and educators.

Carefully designed cross-linguistic child-language studies have also extended and deepened our knowledge of SPECIFIC DOMAINS of linguistic structure and language use. These include issues of sentence processing, spatial reference, and motion events in work referred to earlier by scholars such as Bates and MacWhinney and by Bowerman and Slobin and their associates respectively. Such work is supplemented by in-depth analyses of acquisition of selected linguistic systems in different languages, for example: (a) causative constructions and transitivity alternations (Allen, 1998; Berman, 1993a, 1993b; Bowerman, 1974; Hochberg, 1986; Pye, 1993); (b) passive voice constructions (e.g. Allen & Crago, 1996; Demuth, Jisa, Reilly, 2002; Jakubowicz, 1999; Gelman, 1999). Direction to acquisition of familiar and also comparable child language use is often also evaluated by using naturalistic adult usage in the case of child–adult interactions or semi-structured elicitations and experimental designs. As such, it would appear that child language use is significantly shaped by the social-cognitive context of its use, and this is something that is true of adult language use as well.

As a cross-language comparison study, the Berkeley data base contains information on a large number of languages from around the world. (Evans & Demuth, 2006). These include a wide range of languages from different world language families, as well as a number of non-literate languages, and cover the major world language families. This information is available online at the Cross-Linguistic Data Repository (CLDR) and the Berkeley Language Data Base (BLDB). The data base is currently undergoing a major revision to make it more accessible to researchers and educators.
Cross-linguistic child language research over the past four decades has also made significant contributions to methodology in the domain by application of carefully comparable experimental designs in both production and comprehension, and in elicitation of extended discourse beyond individual lexical items or sentential constructions—for example, by use of the same pictures depicting different types of spatially oriented events, on the one hand, or by eliciting narratives based on the same set of picture-series or a single picture-book story or film. Another important lesson has been use of the same designs both with children at different levels of age-schooling and also with proficiently literate adult speaker–writers as a baseline for comparison of typological preferences. A further direction which should be developed in this respect would be to combine different methodologies focused on the same topic of analysis, as was originally proposed in the Berkeley cross-linguistic study (Slobin, 1982). For example, in order to evaluate children’s knowledge of a certain domain in different languages, it would seem ideal to base structured experimental elicitations on prior analysis of naturalistic adult–child or peer interactions and ethnographic fieldwork, on the one hand, and/or cross-sectional text construction, on the other.

As against the indubitable achievements noted throughout this paper, cross-linguistic comparisons of child language still face complex challenges. Thus, an explicit and generally applicable articulation of what is universal and what particular in first language acquisition (Slobin, 1982) in the interaction between different developmental phases and different ambient languages would greatly advance the field of child language research as a whole. Even if the idea of language universals is abandoned as a myth (Evans & Levinson, 2009), critical questions still remain as to the precise nature and weight of the impact of target language typology on processes of acquisition in different areas of linguistic form and language use. These include questions such as “Which kinds of categories are likely to be most influenced by language, and how this influence manifests itself developmentally?” (Imai & Gentner, 1997) and the relative role and weight of the ‘typological imperatives’ of a given native language (Berman, 1986) impinging on acquisition in terms of shared social–cognitive development.
RUTH A. Berman

While these are formidable demands, it seems reasonable to hope that the critical mass of data now available from diverse languages combined with currently evolving neurobiological and digital technologies could point to interesting new directions for deriving psychologically sound and linguistically well-motivated measures for evaluating the interplay between the variables of developmental phase/ambient language typology/and linguistic systems.

REFERENCES

CROSS-LINGUISTIC COMPARISONS IN CHILD LANGUAGE RESEARCH


