

Continuity in lexical and morphological development in Icelandic and English-speaking 2-year-olds*

ELIN T. THORDARDOTTIR, *McGill University*

SUSAN ELLIS WEISMER }
JULIA L. EVANS } *University of Wisconsin-Madison*

ABSTRACT

Accounts of language development vary in whether they view lexical and grammatical development as being mediated by a single or by separate mechanisms. In a single mechanism account, only one system is required for learning words and extracting grammatical regularity based on similarities among stored items. A strong non-linear relationship between early lexical and grammatical development has been demonstrated in English and, more recently, in Italian supporting a single mechanism view (Caselli, Casadio & Bates 1999, Marchman & Bates 1994). The present study showed a comparable non-linear relationship between vocabulary size and the emergence of verb inflection and sentence complexity in two-year-old speakers of English and Icelandic, a highly inflected language. The study included 96 children within a narrow age range, but varying extensively in language proficiency, demonstrating continuity in lexical and grammatical development among children with typical language development as well as very precocious children and children with expressive language delay. Cross-linguistic differences were noted as well, suggesting that the Icelandic-speaking children required a larger critical mass of vocabulary items before grammatical regularity was detected. This is probably a result of the more complex inflectional system of the Icelandic language compared with English.

Theoretical accounts of language acquisition vary in whether they view

* Address for correspondence: Dr Elin T. Thordardottir, School of Communication Sciences and Disorders, McGill University, 1266 Pine Avenue West, Montreal, Quebec H3G 1A8, Canada. E-mail: elin.thordardottir@mcgill.ca

the lexicon and grammar as being mediated by separate or by common mechanisms. In dual-mechanism views, the emergence of grammar reflects the operation of specific rule-based mechanisms that are dedicated to grammar and whose activation does not depend on lexical development. In contrast, a single mechanism view assumes that the mechanism responsible for learning lexical items is also responsible for extracting grammatical patterns. A key difference in the predictions made by dual mechanism versus single mechanism accounts is in whether the onset of grammar is related systematically to the size of the lexicon. A close relationship between the lexicon and grammatical development has been shown in a number of studies of English (Bates & Goodman 1997, Marchman & Bates 1994), and more recently, Italian (Caselli *et al.* 1999). In the present study, the relationship between lexical growth and the emergence of grammar was examined in Icelandic, a highly inflected language, as well as in English.

Descriptive studies of early language development have shown that it is characterized by remarkable regularity. In spite of differences in children's environments, a developmental sequence has been identified that typically-developing children generally tend to follow. Although the exact sequence is not identical in different languages, cross-linguistic studies have documented regularity in the development of various languages (Slobin 1992). Nevertheless, children vary quite extensively in the rate of language development for reasons that are not well understood. As a result, normally developing children of the same age can vary greatly in their linguistic proficiency. A key question relative to the interdependence of linguistic domains in acquisition concerns whether these domains tend to develop together, or separately with different onset times and different rates of development.

Several previous studies have offered compelling evidence for the existence of a close relationship between linguistic domains in the acquisition of English. These studies have employed large samples of children with normal language development ranging in age from roughly 18 to 30 months and have shown that grammatical development is more strongly related to vocabulary size than to age. In the present study a further examination of this relationship was undertaken by including another language, Icelandic, which differs from English in important structural properties, being highly inflected. In addition, a different approach was used to address age effects. The sample in this study consists of children who were all of the same age, but who varied widely in language proficiency. The inclusion of children of the same age with drastically different vocabulary sizes allows for examination of the relationship between vocabulary and grammar within a narrow

age range; the children were at the age when word combinations are typically starting to appear and grammatical morphology is emerging. Another important feature of the present study is that the sample includes, in addition to children with typical language development, children whose vocabulary size differs significantly from the norm for their age. Children scoring at or below the 10th percentile have a significant delay in expressive language (comprehension skills were not assessed). These children have been referred to as late talkers or children with specific expressive language impairment. The investigation of early vocabulary and morphological development in late talkers is especially interesting given the prominence of morphological deficits demonstrated by English-speaking children with SLI. The sample also included children at the other extreme, with unusually large vocabularies for their age. This group of children has received much less research interest than children with language impairments. However, according to a single mechanism account, the emergence of grammatical morphology should be closely related to vocabulary size for both late developing and precocious children, and this should be true of both the English-speaking and the Icelandic-speaking children.

The most obvious structural difference between Icelandic and English is that Icelandic is highly inflected. Studies have shown that Icelandic children start using grammatical inflections as early as at the one-word stage and subsequently add gradually to the frequency and variety of inflections (Elin Thordardottir & Ellis Weismer 1998). However, full mastery of the inflectional system extends at least into the school years (Ragnarsdottir, Simonsen & Plunkett 1999). As a result of the requirement for grammatical inflections, young Icelandic children have, as a group, a larger mean length of utterance (MLU) than English-speaking age-matched peers, but tend to have smaller vocabularies as measured either by parent report or spontaneous language sampling (Elin Thordardottir & Ellis Weismer, 1996, 1998). These cross-linguistic differences suggest that the relationship between vocabulary and grammar in acquisition is not identical across the two languages. Another source of information on this aspect of highly inflected languages comes from studies of the characteristics of language impairment. Specific language impairment (SLI), defined as deficient language with normal development in other respects, is, in English, characterized by a particular difficulty with grammatical morphology (Leonard 1998). The finding that grammatical morphology can be selectively impaired can be interpreted as evidence for the autonomy of grammatical morphology. Somewhat surprisingly, for children with SLI speaking more highly inflected languages, for example Hebrew,

German, Italian, and Icelandic, grammatical morphology appears not to be as vulnerable as in less highly inflected languages (e.g., Bortolini, Leonard & Caselli 1998, Dromi, Leonard & Shteiman 1993, Elin Thordardottir 2001, Lindner & Johnston 1992). This suggests, again, that the relationship between grammatical morphology and other areas of language development is in some way different in highly inflected languages and in English. Clearly, further study is required to assess whether the close relationship between vocabulary size and grammatical development demonstrated for English is also found in other languages that differ from English in structural properties.

The development of grammatical morphology has traditionally been one of the most obvious manifestations of the existence of rule-based mechanisms in language. Children's first past tense forms are typically high-frequency forms such as 'fell' and 'broke', which are initially produced correctly and are typically assumed to have been learned by rote and are not considered productive. Later on in development, children typically start to make errors in the production of verbs that were previously produced correctly. Thus, 'fell' becomes 'falld', and 'broke' becomes 'brekd'. These overgeneralizations of the regular past tense have provided evidence of the application of a general rule. While the assumption that initial rote learning later becomes productive is widely accepted, as is the assumption that overgeneralization errors are a by-product of this systematicity, accounts vary in how productivity is assumed to come about and in what exactly is meant by rule-based behaviour. A strong dual mechanism view of morphological development has been proposed by Pinker (1991) and articulated further by Marcus and his collaborators (Marcus 1995, Marcus, Brinkman, Clahsen, Wiese & Pinker 1995). According to this account, the use of inflectional morphology is mediated by a specific grammatical rule-based mechanism that is independent of lexical mechanisms. A rule is defined as a mental operation, one that always applies in an all-or-none fashion when the right conditions exist. In this account, the English past tense is formed by a default rule which stipulates that the *-ed* suffix is applied. However, irregular past tense forms are stored in memory and, when they are successfully retrieved, they take precedence over the default rule. Overgeneralizations are the product of over-application of the rule resulting when stored past tense forms fail to block the application of the rule. This occurs because memory can be hampered by performance factors (this can happen to adults under some stressful circumstances, but is especially likely to happen to children). Children's grammars are seen as being of similar structure to that of adult grammars. However, children's memories for irregular forms are not as

good as adults' memories, hence overregularization errors. In a dual mechanism account, mental rules are viewed as a natural property of language, and their development is assumed to be directed by internal factors that govern the development of grammar, and might thus be most strongly associated with age rather than with other linguistic developments. The onset of overgeneralization errors is a direct result of the development of the rule and is unrelated to changes in the composition of the vocabulary (Marcus 1996).

A single mechanism view does not include explicit rules of the kind that are an essential part of the dual mechanism account. Instead, as lexical items are stored, regularities among them are detected and organization takes place. Thus, phonological similarity as well as type and token frequency are among the factors that lead the system to this qualitative shift towards organization (Marchman, Plunkett & Goodman 1997). In order to be able to detect regularities, however, the system must have a minimum number of items to work with. Once this 'critical mass' of lexical items has been acquired, the use of regular patterns such as grammatical inflections increases at an accelerated rate. Importantly, in this type of self-organized system, no explicit rule or dedicated mental operation is required or assumed. Several connectionist simulations have been conducted of past tense acquisition that show the same stages as are evident in children's learning with initial correct use of irregular forms, followed by a period of overgeneralization errors (Plunkett & Marchman 1993, Rumelhart & McClelland 1986).

In spite of the important differences in the underlying assumptions of dual mechanism and single mechanism accounts of morphological development, both accounts offer an explanation of the sequence of events observed in the emergence of grammatical morphology, and many of the predicted outcomes in terms of correct and overgeneralized forms are similar. Both views assume an initial period during which irregular past tense forms are used correctly before overregularization errors appear. In a dual mechanism view, irregular past tenses are being memorized during this time, but the past tense rule has not yet developed (or has not become active). In the single mechanism view, the regular past tense pattern has not been extracted because the system has not acquired a sufficient number of items to detect the regularity. Further, both accounts assume that, once overgeneralization errors appear, there is a period during which some forms are used correctly and some are overgeneralized. In the dual mechanism view, this is because memory of stored items in some cases successfully blocks the application of the all-or-none rule, and in some cases fails to do so because of performance limitations that preclude retrieval of the correct

irregular past tense form. In the single mechanism account, the general pattern that has been extracted does not have an unconditional, all-or-none application, nor is it fully formed from the beginning. Reasons why it may not be applied to certain items include, for example, that the irregular item in question may not bear close phonological resemblance to forms for which regularities have been detected, or the irregular form may be of such high frequency that it is firmly established and is protected from error. The crucial difference between the two accounts is in the relationship of the onset of overregularization errors to the size of the vocabulary. Whereas the dual mechanism account holds that the activation of the rule-based mechanism is mediated by internal factors separate from lexical development, and, therefore, independently of vocabulary size or composition (Marcus 1996), the single mechanism account views morphological regularity as emerging from the lexicon and, hence, as being dependent on it. Therefore, a strong relationship between vocabulary size and use of grammatical inflections is to be expected.

Several previous studies have demonstrated a strong relationship between vocabulary size and grammatical development in English (e.g., Anisfeld, Rosenberg, Hoberman & Gasparini 1998, Bates & Goodman 1997, Fenson, Dale, Reznick, Bates, Thal & Pethick 1994, Marchman & Bates 1994). A number of these studies have employed data obtained from a parent report checklist, the MacArthur Communicative Development Inventory (CDI) (Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly 1993), which includes a vocabulary checklist as well as sections that survey the use of grammatical morphology and syntax. The norming sample of this instrument has permitted examination of large samples of children. The parent report format is an excellent measure of vocabulary size for young children. Since it is based on the parents' observation of their child across a variety of settings, it reflects what the child knows rather than what the child uses in a particular sample (Bates, Bretherton & Snyder 1988). Plunkett & Marchman (1996) have pointed out that it is the set of words the child actually knows which provides the basis for extraction of regularities. In a longitudinal study using both parent report and parent interview data, Bates *et al.* (1988) showed that the measure collected at 20 months which best predicted mean length of utterance (MLU) at 28 months was vocabulary size. This was true even though the assessment at 20 months also included MLU. Marchman & Bates (1994) used the CDI checklist to investigate past tense verb inflection and its relation to vocabulary size. The acquisition of the regular past tense was inferred from overregularization errors. The CDI sections used were Irregular

forms, which lists a number of irregular past tense forms, and Errors, which lists a number of overregularized past tense forms. In addition, a total of 16 verbs occur in the general vocabulary section (where they appear in the bare stem, or infinitive form), as well as in Irregular forms and Errors, which permitted examination of the number of verbs used as bare stems only, versus correctly inflected and overgeneralized. Correct inflection of irregular verbs, as well as overgeneralization errors, were predicted by the size of the verb vocabulary. The use of verb inflection evidenced a nonlinear acceleration once a critical mass of verbs had been acquired, supporting the view that morphological development is continuous with lexical development.

Other studies have used the parent report checklist to examine more global aspects of grammatical development. Fenson *et al.* (1994) examined the relationship between vocabulary size and sentence complexity using two sections of the MacArthur CDI, the vocabulary checklist and the last section, in which parents select from sentence pairs the one that most closely resembles the sentences their child uses. This section is designed to tap syntactic developments occurring in early language development. The results demonstrated a close connection between vocabulary and sentence complexity. The correlation remained strong after age was partialled out. By examining the vocabulary section of the MacArthur CDI, it was shown that the composition of English-speaking children's early vocabularies changes as total vocabulary size increases, with common nouns constituting a progressively smaller percentage of the total vocabulary as the use of predicates (verbs and adjectives) and closed class items increases (Bates, Marchman, Thal, Fenson, Dale, Reznick, Reilly & Hartung 1994). The use of closed class items was related to vocabulary nonlinearly, with a sharp acceleration in growth starting at a vocabulary level of 400 words. The results supported a close relationship between grammatical development and vocabulary size, as well as a critical mass effect in the onset of productive use of closed class items. Recently, Caselli *et al.* (1999) undertook a replication of these findings in Italian using an Italian adaptation of the MacArthur CDI. It was found that Italian children lagged slightly behind age-matched American peers in vocabulary size. However, when children were matched cross-linguistically on vocabulary size, the development of sentence complexity was remarkably similar across languages, replicating the nonlinear pattern demonstrated previously for English. The relative proportions of nouns, predicates and function words were also comparable across languages, though some exceptions were noted, as would be expected from a language as different in structure as Italian is from English.

The study of Anisfeld *et al.* (1998) differs from those just reviewed in that it used direct observational data rather than parent report and it used a longitudinal approach. Five children were observed during play longitudinally to examine the relationship between accelerated vocabulary growth and combinatorial speech. Accelerated vocabulary growth was defined by a deviation from a fitted vocabulary growth curve. It was found that the timing of this growth spurt overlapped with the onset of combinatorial speech. More precisely, combinatorial speech tended to precede the vocabulary spurt. Anisfeld *et al.* concluded that the vocabulary spurt is facilitated by combinatorial speech. This may occur because word combinations may necessitate more specific words than one-word expressions and because the greater syntactic knowledge required to use word combinations may guide vocabulary learning. These findings indicate that the vocabulary spurt is not mediated simply by lexical factors, but rather that lexical and syntactic development work together.

The studies just reviewed all provide evidence of continuity between lexical and grammatical development, supporting a single mechanism view of language acquisition rather than a strict separation of linguistic domains (however, see Marcus 1995, 1996, Marcus *et al.* 1995). It has been proposed that this type of continuity relationship may be a universal of language. Instead of rules being a natural property of languages as proposed by dual mechanism accounts, the nature of language may be the ability to abstract regularities and to self-organize (Bates & Goodman 1997). If this is true, then the traditional view of language as being made up of various separable areas or domains may be in need of serious reconsideration (Bates & Goodman 1997). Some cross-linguistic differences would be expected, based on differences between how languages are structured and how they express meanings, but in each case the emergence of grammar should be related in some lawful way to vocabulary acquisition. Cross-linguistic replication focusing on Italian has shown a comparable relationship between the lexicon and sentence complexity, as was previously demonstrated in English, although some differences were found that were attributed to both structural and pragmatic differences between the languages. The present study is another cross-linguistic examination of the relationship between lexical and grammatical development which focuses on grammatical morphology, an aspect not targeted specifically in the Italian study. Sentence complexity is also examined. Grammatical morphology is particularly interesting in this English/Icelandic comparison because of the much higher degree of complexity of the Icelandic inflectional system. Another unique feature of this study is the inclusion

of children with normal language development as well as children with documented delays in expressive language. The investigation of the connection between lexical and morphological development in late talkers speaking these two languages is especially relevant, given the prominence of morphological deficits in SLI in English and the expectation that such deficits may be less pronounced in SLI in Icelandic. The specific questions asked in this study were whether use of grammatical morphology and sentence complexity is systematically related to vocabulary size in a sample of two-year-old children varying in vocabulary size and whether such a relationship exists in English as well as Icelandic. The examination of grammatical morphology focuses on regular and irregular past tense inflection and associated overregularization errors, following Marchman & Bates (1994). In spite of their many structural differences, English and Icelandic both have verbs that form the past tense in a regular and irregular manner and thus have comparable types of overregularization errors.

METHODS

Participants

A total of 96 children participated in this investigation, 51 Icelandic-speaking (24 girls and 27 boys) and 45 English-speaking (22 girls and 23 boys). All the children were 2 years old, within 2 months (range 22–26 months). The Icelandic children were part of a norming study for the Icelandic parent report employed in this study (Elin Thordardottir 1998); the English-speaking subjects are part of an ongoing study of language development in children with normal language development and late talkers (Ellis Weismer & Evans 2002). The sample size was determined by availability of children in these databases at the time of the study who met the criteria of age and of being reported to have had normal developmental histories (although suspected language delay was allowed). The children span a large range of proficiency, which is representative of the large variability typically observed among normally developing 2-year-olds. Participant characteristics are presented in Table 1, in which the children have been divided into three groups according to language proficiency (high and low extremes, as well as mid-range). These groups were formed *post hoc* in order to demonstrate the proficiency distribution in the sample of children speaking each of the two languages, but children were not selected based on this grouping. Thus, although the sample is constituted of children whose vocabulary scores are distributed on a continuum from low to high, Table 1 demonstrates that, within each language group, most of the

TABLE 1. *Participant characteristics: age and scores on language measures (broken down by proficiency level based on vocabulary size) – means and (standard deviations)*

Icelandic-speaking children						
	> 90% ile (n = 12)		11–89% ile (n = 24)		< 10% ile (n = 15)	
Percentile	90	(0)	55	(16.74)	10	(0)
Age	24.0	(1.41)	24.0	(1.23)	23.8	(1.52)
Total words	519	(128)	295	(95)	52	(20)
Total verbs	96	(25)	60	(23)	4	(3)
UTT3	9.13	(4.10)	6.46	(2.98)	2.00	(1.37)
Word comb.	100%		91.3%		46.7%	

English-speaking children						
	> 90% ile (n = 8)		13–89% ile (n = 27)		< 12% ile (n = 10)	
Percentile	90.12	(3.6)	51.3	(21.33)	6.1	(4.6)
Age	25.1	(1.35)	23.9	(1.15)	23.9	(1.28)
Total words	572	(38)	316	(117)	75	(45)
Total verbs	94	(4)	45	(26)	5	(6)
UTT3	7.83	(3.05)	4.27	(1.75)	2.13	(1.03)
Word comb.	100%		92.6%		60.0%	

Percentile is based on the Total Vocabulary section of the parent report checklist (Section IIA). Note that for the Icelandic data, the standard deviation in the percentile ranks for some of the groups appears as 0. This is because the norm database does not permit percentile ranks to be determined with greater precision even though the raw scores of children within the groups vary.

UTT3 is the mean length in morphemes of the child's three longest utterances as reported by the parent.

Word comb. is the percentage of children in the group reported to have started combining words.

children fell within the middle range, and a smaller number obtained extreme scores. It should be noted that, due to the smaller sample size in the Icelandic parent report norms than in the English norms, percentile ranks cannot be determined as accurately for Icelandic as for English. Therefore, the Icelandic children in the extreme groups (above the 90th and below the 10th percentiles) cannot be differentiated more

precisely in terms of percentile rank, although their raw scores differed. Therefore, the standard deviation for these groups appears as 0. This should not be interpreted as no variability existing within these groups. Table 1 presents, in addition, several other parent report scores for the children, including the number of verbs in their productive vocabulary (Total verbs), the mean length of their three longest utterances (UTT3), which gives an indication of syntactic proficiency, and the proportion of children in each group reported to have started producing two-word combinations (Word comb.).

All the children were monolingual speakers of their native language (Icelandic or English). None of these children was reported to have had major illnesses or a developmental history that raised concerns about their health or general development. The English-speaking children included 43 white children, one biracial African-American/white child, and one child for whom ethnic background was not reported. All 45 English-speaking children came from two-parent homes. Nineteen of these children were firstborns and 10 were only children. The mean number of years of maternal education for this group was 16.4 (SD 1.92). The Icelandic children were an ethnically homogeneous group; 39 children came from two-parent homes, 6 from single-parent homes, and 6 did not report family structure. Nineteen of the Icelandic children were firstborns and 19 were only children. The mean number of years of maternal education was 14.07 (SD 3.55). No attempt was made to achieve a direct match of vocabulary scores cross-linguistically. Rather, the language groups have a similar distribution of percentile ranks which were determined for the children based on a comparison with same-language peers. In fact, data in Table 1 indicate that raw vocabulary scores corresponding to middle and extreme percentiles are not identical across the languages, with Icelandic children obtaining somewhat lower vocabulary scores than the English-speaking children, consistent with previous studies (Elin Thordardottir & Ellis Weismer 1996).

Procedure

Data were collected by parent report checklists, the MacArthur Communicative Development Inventory (CDI) for Toddlers (Fenson *et al.*, 1993) and its Icelandic adaptation (Elin Thordardottir & Ellis Weismer 1996, Elin Thordardottir 1998). The MacArthur CDI contains two main parts: a vocabulary checklist and a grammatical checklist with sections addressing various aspects of grammatical morphology and syntax. The Icelandic adaptation is analogous in form, but the content was modified to reflect characteristics of the Icelandic language. The forms are

sufficiently similar to allow meaningful cross-linguistic comparisons, at the same time taking into account inherent cross-linguistic differences. The measures used in this study and the sections of the parent report checklists from which they were derived were as follows: (1) vocabulary size, or total number of checked words on the vocabulary checklist; (2) number of verbs, or total number of checked items in the verb subsection of the vocabulary checklist; (3) number of different regular noun suffixes used (Section II-A, 'Word Endings, Part 1'); (4) number of different verb suffixes used (Section II-A, 'Word Endings, Part 1'); (5) number of correct irregular verb forms used (Section II-B, 'Word Forms'); (6) number of overregularized past tense verb forms used (Section II-C, 'Word Endings, Part 2'); (7) syntactic complexity score (Section II-E, 'Complexity'). All these measures reflect number of types, not tokens. Each section of the parent report contains a number of items, from which the parents select the ones their child has used at least once. These measures allowed analysis of the development of regular inflections, correct use of irregular forms and subsequent overregularization of these forms as a function of vocabulary size, as well as the relationship of the development of sentence structure to vocabulary size.

The Icelandic version of the parent report checklist was modified to reflect adequately characteristics of the Icelandic language (both cultural and pragmatic differences affecting the composition of early vocabularies, and structural differences, Icelandic being a highly inflected language; see detailed description in Elin Thordardottir & Ellis Weismer 1996). As a result, the contents of some of the subsections differ between languages, in number of items in some sections and in the grammatical structures surveyed. The Icelandic vocabulary checklist is longer (705 versus 680 items in the English version). The verb section differs somewhat between the checklists. In the English version, lexical verbs appear in the section 'Action words' (103 items) and grammatical verbs in an additional section, 'Helping verbs' (21 items). In contrast, the Icelandic checklist has a larger verb section (121) which contains both lexical and grammatical verbs. For comparability between the languages, total verb vocabulary was determined in Icelandic by the score obtained on the single verb section (121 items), and in English, by adding action words and helping verbs (124 items).

There are differences in the inflectional suffixes surveyed in Section II-A. In both checklists this section presents the suffixes, gives examples of their use and asks the parents to indicate whether their child uses the suffix 'never', 'sometimes' or 'often'. Credit is given if the parents check one of the two latter choices. The regular noun suffixes include the

plural *s*, and the possessive *s* in the English version. In the Icelandic checklist the inflections surveyed are the plural and the genitive case. These inflections are largely analogous in function to the English noun suffixes, but their form varies greatly across nouns and is not further specified in the checklist. A third noun suffix is included in the Icelandic checklist: the definite article. Although the article is a bound morpheme in Icelandic, it is not an inflectional morpheme (but must itself be inflected for case and number similar to a noun). It was not included for the purposes of the present cross-linguistic comparison. Regular verb suffixes surveyed on the English form included the regular past (*-ed* suffix or allomorph) and the present progressive form (*-ing* suffix). The Icelandic regular verb forms were the imperative mood (*-ðu, -iði* suffix or allomorph), the regular past tense (*-aði* suffix or allomorph) and the past participle. The total number of verb inflections surveyed was thus two in English and three in Icelandic. The Icelandic checklist contains, in addition, several compound tenses which were not included in the present analysis because they involve use of auxiliaries rather than inflectional morphology.

The section of irregular verb forms (Section II-B) on the English checklist includes 20 items, whereas the Icelandic version includes 17. The English section on overregularized forms (Section II-C) includes 16 verbs whereas the Icelandic version includes 18 verbs. In each case, more than one way of overregularizing a verb may be included on the checklist, e.g., the past tense of 'eat' may become 'eated' or 'ated', and the past tense of the Icelandic verb '*leika*' may become '*leikaði*' or '*leikti*'. The English checklist allows for either possibility to be checked, whereas the Icelandic version lists both possibilities together. For comparability across languages, if either possibility was checked in the English checklist, this was counted as one instance of overgeneralization, involving one verb. Only one instance was encountered in which two types of overgeneralization were checked for a single verb. Another difference between the two versions is that the English overgeneralized verb forms are all past tense forms. In contrast, only 13 of the 18 items on the Icelandic list are past tense forms. Three of the remaining five are incorrectly formed past participle forms that are frequently encountered in children's language production, and two are overgeneralized person endings. Finally, the last Section (II-E) involves sentence pairs from which the parents select the one that most resembles the kinds of sentences used by their child. The English checklist includes 37 sentence pairs, whereas the Icelandic checklist includes 42 pairs. It should be re-emphasized that this study involved analysis of relationships between vocabulary and other aspects of

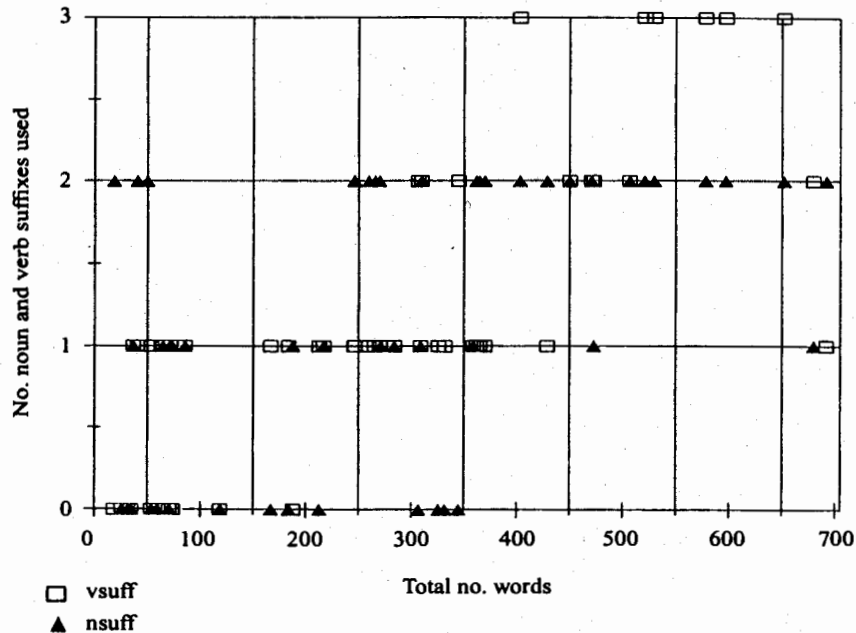


Fig. 1. Number of different types of regular noun and verb inflectional suffixes (nsuff, vsuff) used by the Icelandic-speaking children as a function of Total number of words; maximum score is 2 for noun suffixes and 3 for verb suffixes.

(Note: Each child in the sample is represented by two data points, one showing the number of different noun suffixes used and one showing the number of types of verb suffixes used; these points are situated on the x-axis at the point corresponding to the child's vocabulary size.)

language (morphology and syntax) within each language rather than direct cross-linguistic comparisons of individual measures.

RESULTS

Figures 1 and 2 show the children's use of noun and verb suffixes as a function of vocabulary size. For both Icelandic- and English-speaking children, the reported use of both types of suffixes increased gradually with increasing size of the vocabulary. However, the vocabulary size

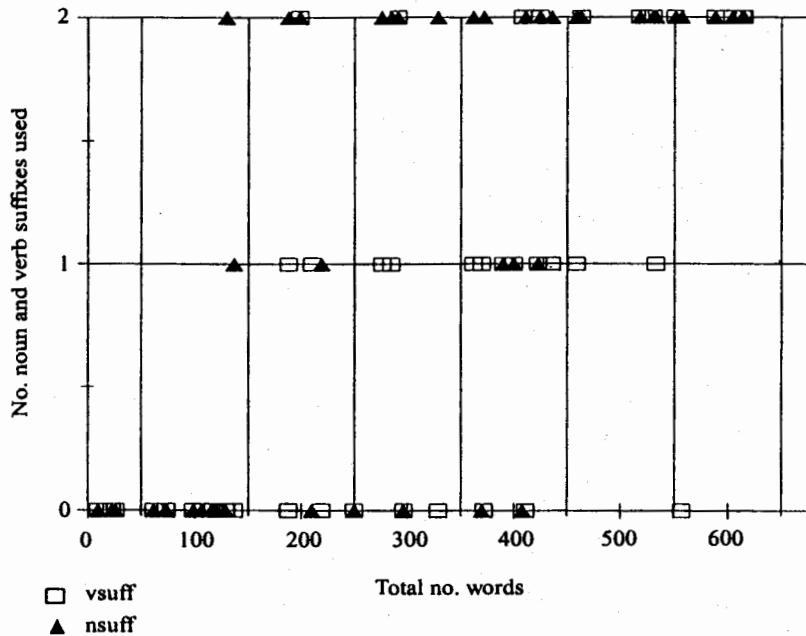
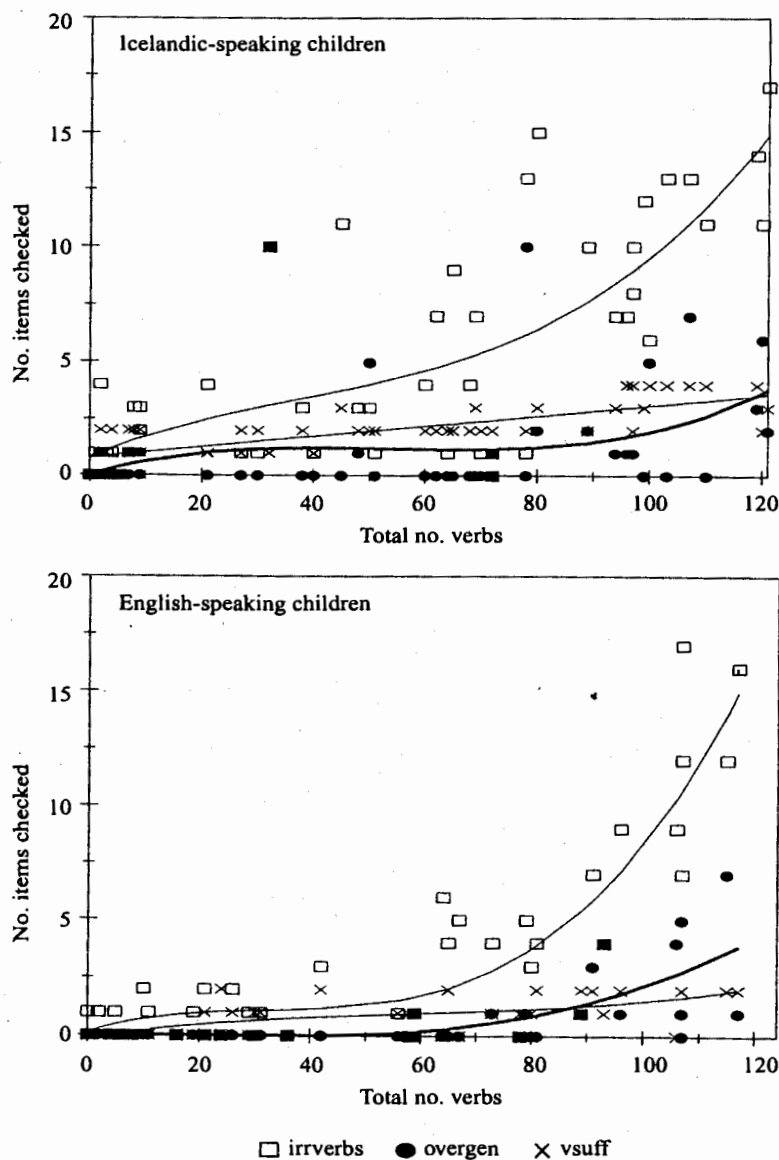


Fig. 2. Number of different types of regular noun and verb inflectional suffixes (nsuff, vsuff) used by the English-speaking children as a function of Total number of words; maximum score for each of the noun and verb suffixes is 2. (See Note for Fig. 1.)

associated with first use of inflectional suffixes varied across languages, with 130 words for English-speaking children, but as few as 19 words for the Icelandic-speaking children. English-speaking children with vocabularies of less than 130 words were found to use no inflectional suffixes, whereas children with vocabularies exceeding 550 words used all the suffixes that were surveyed, two types of noun inflection and two types of verb inflection. Children with vocabularies between these extremes were at various stages in the acquisition of the inflectional suffixes. The Icelandic children started their use of suffixes earlier than the English-speaking children. However, at the maximum vocabulary of 705 words in the Icelandic checklist, some of the children were still not using all the inflections that were surveyed. While some of the Icelandic children were reported to have started using both noun and

verb inflections as early as at a vocabulary size of 19 words, many started only later. However, at a vocabulary size of roughly 200 words, all the Icelandic children had started using some verb inflections, and by 350 words all had started inflecting nouns as well. At the maximum vocabulary size of 705 words, all the Icelandic-speaking children were using both noun suffixes, but some were still using only one or two of the three verb suffixes that were surveyed. The use of the regular inflectional suffixes was significantly and strongly correlated with vocabulary size (and with the size of the verb lexicon) in both languages (Icelandic, $r = 0.749$, English, $r = 0.787$, both correlations significant at the 0.01 level).

The results of Sections II-B and II-C, verb inflections, including correctly used irregular verb forms and overgeneralized verb forms, are plotted in Figs 3 (Icelandic children) and 4 (English-speaking children) as a function of size of the verb vocabulary. It should be added that the size of the verb vocabulary is highly correlated with total vocabulary (Icelandic: $r = 0.968$, English: $r = 0.982$, both significant at the 0.01 level) and that similar plots were, therefore, obtained whether the x -axis shows verbs or total vocabulary. Figures 3 and 4 also show the use of regular verb suffixes from Section II-A. The data for each of the three variables were fitted using Chebyshev polynomials (cf. Cheney & Kincaid 1985). In both languages the use of both correct irregular forms and overgeneralization errors clearly increased with the size of the verb lexicon. Moreover, the pattern of growth in both languages showed evidence of a non-linear acceleration related to the size of the verb lexicon. Deviation from linearity was evaluated by goodness-of-fit testing, applying linear and quadratic models. The interest was in whether the data were accounted for by a curvilinear model, providing evidence of a nonlinear acceleration in growth of grammatical variables as a function of vocabulary size. For the English data, both the irregular verb forms curve and overgeneralizations curve were accounted for better by the quadratic model than by the linear model (irregular verbs: quadratic: $r^2 = 0.700$, $p = 0.000$, linear: $r^2 = 0.547$, $p = 0.000$; overgeneralizations: quadratic: $r^2 = 0.519$, $p = 0.000$, linear: $r^2 = 0.349$, $p = 0.000$). The location of the nonlinear acceleration in growth was evaluated using log-log plotting of the data, revealing that it occurred at a verb vocabulary of approximately 55 verbs for both measures. The Icelandic irregular forms curve similarly fitted a quadratic model better than a linear model (quadratic: $r^2 = 0.617$, $p = 0.000$; linear: $r^2 = 0.584$, $p = 0.000$). However, the relationship between overgeneralization errors and vocabulary in Icelandic was less strong than in English, and the quadratic fit was not statistically significant (quadratic: $r^2 = 0.114$,



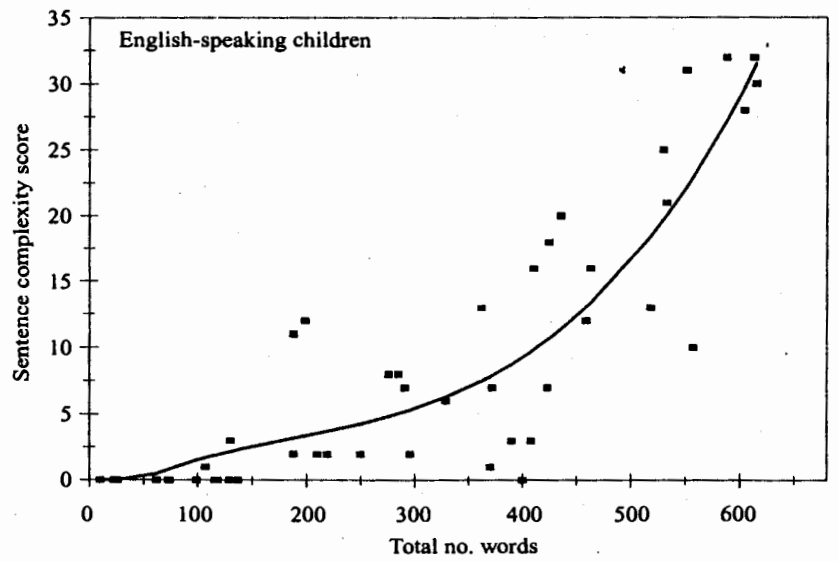
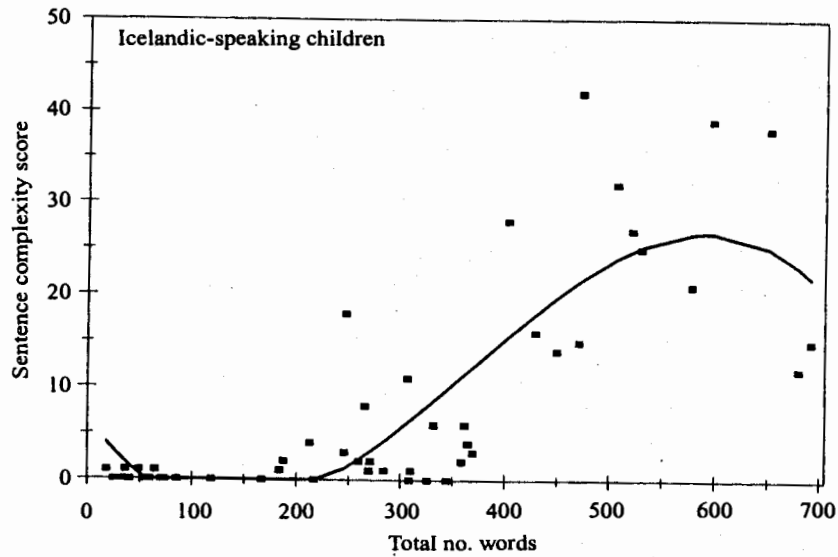
Figs 3 and 4. Individual data points and fitted curves showing the number of irregular verb types used correctly (irrvverbs), number of irregular verb types used in an overgeneralized regular form (overgen), and number of types of regular verb inflections (vsuff) used by the Icelandic- (Fig. 3) and English-speaking children (Fig. 4) as a function of Total number of verbs. The curve for overgeneralization errors is in bold.

$p = 0.059$; linear: $r^2 = 0.109$, $p = 0.019$). Inspection of the overgeneralization data points reveals large variability, especially at low verb vocabularies, with three children with low vocabularies using an unusual number of overgeneralizations compared with the rest of the sample. The goodness-of-fit test was repeated with the overgeneralization scores of these three children removed. The resulting curve fitted the quadratic model better than the linear model (quadratic: $r^2 = 0.423$, $p = 0.000$; linear: $r^2 = 0.269$, $p = 0.000$). Log-log plotting of the Icelandic data indicated that a marked acceleration in growth of irregular forms occurred at a verb vocabulary of approximately 70 verbs and of overgeneralizations at approximately 80 verbs.

The development of sentence complexity (Section II-E) is shown in Figs 5 and 6 for the Icelandic- and English-speaking children, respectively. For the English-speaking children, sentence complexity increased in a curvilinear fashion as a function of vocabulary size which was accounted for somewhat better by a quadratic than a linear fit (quadratic: $r^2 = 0.684$, $p = 0.000$; linear: $r^2 = 0.628$, $p = 0.000$). The sentence complexity pattern for the Icelandic-speaking children is overall consistent with that of the English-speaking children in that children with larger vocabularies in general used more complex sentences. Whereas children with 200 words and less in their vocabularies obtained a maximum sentence complexity score of 2, a clear upward turn occurred at a vocabulary size just exceeding 200 words and continued up to a vocabulary of 600 words. However, at vocabulary sizes exceeding 600 words, an unexpected downward turn was seen. The pattern of sentence complexity for the Icelandic children was best captured by a cubic fit (cubic: $r^2 = 0.678$, $p = 0.000$; quadratic: $r^2 = 0.576$, $p = 0.000$).

DISCUSSION

The results of this study provide evidence of a strong and systematic relationship between lexical and grammatical development in both English and Icelandic. Among children of the same age varying in linguistic proficiency, the emergence of inflectional morphology and the development of sentence structure were significantly related to vocabulary size. The grammatical morphology results for the English-speaking children replicate earlier findings of Marchman & Bates (1994), showing evidence that both correct irregular verb forms and overgeneralized verb forms were clearly related to vocabulary size in a systematic nonlinear fashion with relatively little variability. In addition, the present results extend these findings to children with atypical language development, including children with expressive language deficits and precocious



Figs 5 and 6. Individual data points and fitted curve for sentence complexity score (Section II-E of parent report checklist) as a function of Total number of words for the Icelandic- (Fig. 5) and English-speaking children (Fig. 6)

children. This is an important addition, demonstrating that among children of the same age grammatical development is closely related to vocabulary size. Within this age range, children with atypically small vocabularies were similarly deficient in grammatical development, and children with very precocious vocabulary development were similarly advanced in grammatical development.

As expected based on previous work, neither regular inflections nor overregularized forms were used until a certain number of vocabulary items were acquired, supporting the claim that a critical mass of words is required for the system to extract regularities. Furthermore, examination of the data (Fig. 4) revealed that the order of appearance of verb forms was as expected: initially, children used primarily uninflected verbs. Some limited use of correct irregular forms occurred at low vocabulary levels, but increased markedly at higher vocabulary levels. Overgeneralized forms, however, were used only by children with larger vocabularies. Further, it was found that all children who used overgeneralized forms were reported to use regular verb inflections, lending further support to the claim that overgeneralization errors reflect productive use. The development of sentence complexity demonstrated a similar nonlinear pattern as a function of vocabulary size and replicated previous findings (Fenson *et al.* 1994). The data for the English-speaking children thus offered clear support for a strong relationship between lexical and morphological development.

The Icelandic children showed similar overall patterns to the English-speaking children, as well as important differences. The emergence of regular verb and noun inflections showed a clear relationship to vocabulary size. In both languages, children with small vocabularies used no inflectional morphology and all children with large vocabularies did. Moreover, the larger the vocabulary, the greater the variety of suffixes tended to be. However, Icelandic children started using regular inflectional morphology earlier than did the English-speaking children, some at vocabulary sizes as low as 19 words, compared with 130 for the English-speaking children. The groups also differed at the upper end of the vocabulary distribution. Whereas all the English-speaking children with vocabularies exceeding 400 words used all the regular noun and verb inflections surveyed by the parent report form, some of the Icelandic-speaking children producing close to 700 words were still lacking some noun and/or verb inflections, indicating a more protracted course of morphological development. It should be remembered that the Icelandic-speaking children also have more morphemes to learn. The very early use of grammatical inflections in Icelandic is consistent with previous reports (e.g., Elin Thordardottir & Ellis Weismer 1998)

which found that children use certain types of inflectional morphology consistently and correctly in language samples with MLUs barely exceeding 1. The protracted course of development is also consistent with previous studies (e.g., Ragnarsdottir *et al.* 1999).

The Icelandic children's use of correct irregular as well as overregularized verb forms was also systematically related to vocabulary size. As for the English-speaking children, this relationship demonstrated a nonlinear break, and showed the same general order of appearance of uninflected stems, correct irregulars, and regular inflections and overregularized forms. However, even cursory inspection of Figs 3 and 4 reveals important cross-linguistic differences. The most obvious difference is that there is much greater variability in the data of the Icelandic children. Although the use of correct irregular forms definitely increased systematically with growing vocabulary size, this relationship was not as tight as it was in the English data. The variability was even more pronounced for overgeneralizations. In fact, as reported in the results, a quadratic model fitted to the overgeneralization data was not statistically significant. Examination of individual data points and the fitted curve indicated that most of the Icelandic-speaking children with vocabularies of less than 80 verbs used very few or no overgeneralized verb forms. At vocabularies exceeding 80 verbs, however, the use of such forms increased markedly for the group in a non-linear fashion. Three children with vocabularies of less than 80 verbs were clear exceptions to this general pattern, using as many as 10 overgeneralized forms; the most reported for any of the Icelandic children. Given the strength of the general pattern of overgeneralizations (involving 48 of the 51 children) and its consistency with the use of correct irregular forms, it was concluded that the high use of over-generalization of the three above-mentioned children was atypical and non-representative. When these three children were removed from the sample, a significant nonlinear relationship was evident between overgeneralizations and verb vocabulary size.

The removal of these three atypical children based on their not fitting the general pattern still leaves unanswered questions. Why was atypical behaviour of this kind observed in the Icelandic sample only? How did these children arrive at their high use of overgeneralized forms so early in their development? One possible answer is that parent report did not accurately represent the behaviour of these children. It can be speculated that the high degree of complexity of the language made it hard for parents to report accurately on the inflectional use of their children. It is unlikely, however, that this is a factor in most cases. First, the reliability of both the vocabulary and grammatical sections of the Icelandic parent

report checklist has been verified by comparison with language sample data and found to be very high, ranging from 0.89 to 0.93 for the vocabulary list and from 0.81 to 0.89 for Section II-A, Word Endings (Elin Thordardottir & Ellis Weismer 1996). Secondly, the regularity of the data for most of the children in this study supports reliability in reporting of grammatical morphology. Nevertheless, the high reliability of the checklist does not preclude the possibility of significant error in reporting in some number of cases. Further investigation using direct measures such as language samples is required to clarify the extent of variability that can be expected in a highly inflected language such as Icelandic, as well as the extent to which individual differences exist.

Another important cross-linguistic difference in grammatical morphology was in the location on the x -axis at which the non-linear break occurred in the use of correct irregulars and overgeneralized forms: at approximately 55 verbs for the English-speaking children, but not until 70 to 80 verbs for the Icelandic children. This indicates that the Icelandic-speaking children required a larger critical mass of verbs to be able to detect the regularity that led them to overregularization. This occurs in spite of the fact that the initial use of regular inflections starts at a much lower vocabulary size for the Icelandic- than the English-speaking children. A possible interpretation of these findings that takes into account the complexity of the Icelandic inflectional system is as follows. The early use of inflectional morphology is not surprising in a language in which the majority of words must be inflected. The input that children receive reflects this fact. However, it is likely that the earliest use of inflectional morphology is not productive, but instead involves the use of the inflected forms learned by rote and associated initially with the contexts in which they were learned, resulting in mostly correct use.

Previous research has indicated that Icelandic children do use a number of inflected forms very early on in spontaneous language and that they make relatively few errors at first (Elin Thordardottir & Ellis Weismer 1998). Overgeneralized forms, however, indicate productive use. It may be that the greater complexity of the Icelandic inflections, with each word assuming a number of different forms depending on the context, requires a greater number of instances to allow extraction of a general pattern and, therefore, a larger critical mass of stored vocabulary items before the onset of productive use of grammatical morphology. A similar cross-linguistic delay in morphological acquisition was shown recently by Ragnarsdottir *et al.* (1999) in a study comparing error patterns of children aged 4, 6 and 8 in a past-tense elicitation task in Norwegian and Icelandic. Overall, findings of the study supported an

exemplar-based account of morphological acquisition, one in which general patterns are inferred from similarities between previously learned items. At age 4, Icelandic-speaking children lagged behind their Norwegian counterparts in the correct use of the past tense of strong (irregular) verbs, but caught up with them by age 6. As Ragnarsdottir *et al.* pointed out, this cross-linguistic difference may be related to the fact that the morphological system of Icelandic is considerably more complex than the Norwegian system. It may be that, although children speaking the two languages use similar processes in their acquisition of the inflectional system, those learning the more highly inflected language need more instances to arrive at correct use.

The relationship of sentence complexity and vocabulary size also showed cross-linguistic similarities as well as differences. The English-speaking children showed evidence of a gradual increase in sentence complexity scores, starting at low vocabulary levels with an acceleration occurring at a vocabulary size of approximately 400 words. In contrast, the Icelandic children evidenced a nonlinear break at a lower vocabulary size than the English-speaking children, with a little over 200 words. The downward turn at vocabulary sizes exceeding 600 words is unexpected and results from two participants with unusually low sentence complexity scores in comparison with their reported vocabulary size. These two children both fall within normal limits compared to age norms in their sentence complexity, one scoring between the 25th and 50th percentiles and the other between the 50th and 75th. However, both have vocabulary scores that place them above the 90th percentile. Further examination of their forms was undertaken to evaluate whether their sentence complexity score was representative of their ability. Another section of the checklists, Section II-D, asks parents to give examples of their children's longest utterances. For one of these two children, the score on this section is commensurate with Section II-E, sentence complexity. In the other case, the parents did not give examples, stating that the child talked incessantly and that they were unable to give representative examples. This suggests that the sentence complexity score underestimated this child's ability. In conclusion, it appears, as was the case with grammatical morphology, that syntactic development is in general systematically related to vocabulary size for most children. However, there is more variability in the data of the Icelandic-speaking children, the source of which will require further research.

The findings of this study provide further cross-linguistic support for the interdependence of grammar and the lexicon in acquisition in a language that differs importantly from English in its structure.

Consistent with previous research on Italian (Caselli *et al.* 1999), it was also found that, although a relationship was evident between vocabulary size and morphological and syntactic development in the two languages, this relationship is not identical across the languages, reflecting differences in linguistic structure and possibly cultural differences in language use. Thus, the data provided evidence that the size of the critical mass required for children to extract the regular past tense inflection differs across the languages, being somewhat larger in Icelandic than in English. The present investigation of noun inflections also uncovered cross-linguistic differences. However, the examination of noun inflections did not differentiate between nonproductive and productive use, as was the case for past tense. Further studies focusing in greater detail on other aspects of grammatical development are required to examine the effect of vocabulary development on other types of grammatical morphemes.

In addition to extending previous findings cross-linguistically, this study demonstrates that the interdependence of vocabulary and grammar holds for children with average to superior language skills, as well as for children with significant expressive language delays – a group of children often referred to as late talkers. Research has shown that a proportion of late talkers are later identified as having specific language impairment. The observation that the relationship between early vocabulary development and morphological development is similar across children of dramatically different levels of proficiency, including late talkers, suggests that a persistent difficulty in the area of morphology may originate from problems in a broader area than just morphology and may include, for example, vocabulary size or pattern extraction abilities. The late talkers included in this study had such low vocabulary sizes that very limited use of morphology could be expected. Further research is needed to investigate early vocabulary-grammar relationships demonstrated by this group of children once their vocabularies reach the size where a non-linear acceleration would be expected to occur.

The strong relationship between vocabulary size and morphological and syntactic development has been interpreted here as being consistent with the single mechanism account, but not with the dual mechanism account, based on the fact that the dual mechanism account does not predict that a significant relationship exists between lexical and grammatical domains in acquisition. A further assumption, in accordance with the single mechanism account, is that the relationship between the domains is such that grammatical inflection emerges from stored vocabulary. However, the correlational data offered in this study and a number of previous studies does not prove the direction of the

relationship. It has been suggested that a relationship in the opposite direction may exist between lexical and syntactic development. In a recent study, Anisfeld *et al.* (1998) concluded from observational data collected longitudinally from five children that the emergence of combinatorial speech co-occurred with accelerated vocabulary learning. However, the timing of development in the two areas suggested that the onset of combinatorial speech tended to precede any observed vocabulary growth spurt, indicating that the effect was such that the use of word combinations facilitated the learning of new vocabulary. The study of Anisfeld *et al.* differs in important ways from the present study, precluding direct comparisons. In addition to being cross-sectional, the present study examined vocabulary development in terms of cumulative size rather than rate of development. Both studies, however, as well as other previous work, clearly indicate that a strong relationship exists between lexical and grammatical development. Further research is required to establish the nature and direction of this relationship which is likely to prove to be bidirectional and interactive.

Finally, another general point of this study is that it highlights the large variability in language proficiency that exists between children of the same age. This underscores the need to take language proficiency into account when participants are matched for research purposes. In many cross-linguistic studies, participants have been matched across languages on age alone, perhaps largely due to difficulty in cross-linguistic comparison of language proficiency. In future cross-linguistic investigations, it would be beneficial to assess children in each language group relative to appropriate normative data in their own language (see Elin Thordardottir & Ellis Weismer 1998), since clearly, age matching does not guarantee that the children are at a comparable level of development.

REFERENCES

- Anisfeld, M., Rosenberg, E., Hoberman, M. & Gasparini, D. (1998). Lexical acceleration coincides with the onset of combinatorial speech. *First Language*, **18**, 165–84.
- Bates, E., Bretherton, I. & Snyder, L. (1988). *From First Words to Grammar* (Cambridge: Cambridge University Press).
- Bates, E. & Goodman, J. (1997). On the inseparability of grammar and the lexicon: evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, **12** (5/6), 507–84.
- Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, S., Reilly, J. & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, **21**, 85–123.
- Bortolini, U., Leonard, L. & Caselli, M. (1998). Specific language impairment in

- Italian and English: evaluating alternative accounts to grammatical deficits. *Language and Cognitive Processes*, 13 (1), 1–20.
- Caselli, C., Casadio, P. & Bates, E. (1999). A comparison of the transition from first words to grammar in Italian'. *Journal of Child Language*, 26, 69–111.
- Cheney, W. & Kincaid, D. (1985). *Numerical Mathematics and Computing*, 2nd edn (Monterey, CA: Brook/Cole Publishing Company).
- Dromi, E., Leonard, L. & Shteiman, M. (1993). The grammatical morphology of Hebrew-speaking children with specific language impairment: some competing hypotheses. *Journal of Speech and Hearing Research*, 36, 760–71.
- Elin Thordardottir (2001). Morphological errors by Icelandic children with specific language impairment: a question of information value? Poster presented at the Symposium on Research in Child Language Disorders, June 2001, Madison, Wisconsin.
- (1998). *Orðaskil: Málþroskapróf* (Reykjavik, Iceland: Framsaga).
- Elin Thordardottir & Ellis Weismer, S. (1998). Mean length of utterance and other language sample measures in early Icelandic. *First Language*, 18, 1–32.
- (1996). Language assessment via parent report: development of a screening instrument for Icelandic children. *First Language*, 16, 266–85.
- Ellis Weismer, S. & Evans, J. (2002). The role of processing limitations in early identification of specific language impairment. *Topics in Language Disorders*, 22, 15–29.
- Fenson, L., Dale, P., Reznick, S., Bates, E., Thal, D. & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59 (serial no. 242).
- Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. (1993). *Technical Manual for the MacArthur Communicative Development Inventories* (San Diego, CA: San Diego State University).
- Leonard, L. (1998). *Children with Specific Language Impairment* (Cambridge, MA: MIT Press).
- Lindner, K. & Johnston, J. (1992). Grammatical morphology in language impaired children acquiring English or German as their first language: a functional perspective. *Applied Psycholinguistics*, 13, 115–29.
- Marchman, V. & Bates, E. (1994). Continuity in lexical and morphological development: a test of the critical mass hypothesis. *Journal of Child Language*, 21, 339–66.
- Marchman, V., Plunkett, K. & Goodman, J. (1997). Overregularization in English plural and past tense inflectional morphology: a response to Marcus (1995). *Journal of Child Language*, 24, 767–79.
- Marcus, G. (1996). Why do children say 'brokek'? *Current Directions in Psychological Science*, 5 (3), 81–5.
- (1995). Children's overregularization of English plurals: a quantitative analysis. *Journal of Child Language*, 22, 447–59.
- Marcus, G., Brinkman, U., Clahsen, H., Wiese, R. & Pinker, S. (1995). German inflection: the exception that proves the rule. *Cognitive Psychology*, 29, 189–256.
- Pinker, S. (1991). Rules of language. *Science*, 253, 530–5.
- Plunkett, K. & Marchman, V. (1996). Learning from a connectionist model of the acquisition of the English past tense. *Cognition*, 61, 299–308.
- (1993). From role learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition*, 48, 21–69.
- Ragnarsdottir, H., Simonsen, H. & Plunkett, K. (1999). The acquisition of past tense morphology in Icelandic and Norwegian children: an experimental study. *Journal of Child Language*, 26, 577–618.
- Rumelhart, D. & McClelland, J. (1986). On learning the past tense of English verbs. In J. McClelland & D. Rumelhart (eds), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Cambridge, MA: MIT Press).
- Slobin, D. (1992). *The Crosslinguistic Study of Language Acquisition*, Vol. 3 (Hillsdale, NJ: Lawrence Erlbaum Associates).