

19 Morphology in Language Acquisition

EVE V. CLARK

1 Language acquisition

Children typically begin to say their first words between twelve and twenty months of age. And they produce systematic morphological modulations of those words within their first year of talking. As they move to more complex expression of their meanings, they add grammatical morphemes – prefixes, suffixes, prepositions, postpositions, and clitics. On nouns, for example, they start to add morphemes to mark such distinctions as gender, number, and case; on verbs, they add markers for aspect, tense, gender, number, and person. Within a particular language, children's mastery of such paradigms may take several years. There are at least three reasons for this: (a) some meaning distinctions appear to be more complex conceptually than others, and so take longer to learn; (b) some paradigms are less regular than others, and they too take longer to learn; and (c) language typology may affect the process of morphological acquisition: suffixes, for instance, are acquired more readily, and earlier, than prefixes.

In order to acquire noun and verb morphology, children must first analyze the structure of words heard in input, identify stems and affixes, map consistent meanings onto them, and then begin to use those stems and affixes in new combinations. This process of analyzing form and assigning meaning is a prerequisite for the acquisition of inflectional morphology. It is also a prerequisite in the acquisition of word formation. Children begin to use some word-formation processes at around the same time as their first inflections. In particular, they produce novel compounds formed from simple stem combinations (often called *root* compounds). Next, during their second year of speech, as some inflectional paradigms become established, they also begin to produce a few derivational affixes in novel word forms. These emerge in greater numbers between ages three and four, in both derived and compound innovations.

1.1 Some issues in acquisition

The acquisition of morphology in inflection and word formation raises a number of questions about both morphology and the process of acquisition. Languages differ in the extent to which they rely on inflectional morphology to mark grammatical distinctions and grammatical relations. In essence, languages range from analytic (with virtually no inflectional morphology, as in Chinese) to synthetic (with fairly extensive reliance on inflections, some for grammatical relations and agreement, and many marking several distinctions at once, as in Spanish), to agglutinative (with highly regular and systematic inflections, each marking a separate distinction, as in Turkish). In word formation, there is just as much variety, with some languages relying almost exclusively on compounding to form new words, others relying mainly on derivation, and many others relying on some mix of the two. Derivation may include both affixal and zero-derived forms, and some affixes commonly appear in compounds too. One issue for acquisition, then, is the extent to which language typology affects the process of learning: Do particular typologies help or hinder?

In acquisition, the domain of morphology is the word. Inflectional affixes are added to words or stems to form words. Each one must be identified and assigned some meaning. But inflections also mark grammatical relations through agreement within or across phrases, so the domain of an inflection may go beyond the word. Take number. Children acquiring English or Spanish must learn that nouns denoting more than one instance of an entity are used in a plural form, with the addition of the relevant inflection. In Spanish, though not in English, any adjective that accompanies that plural noun must also be marked for plural. In addition, if that noun happens to be the grammatical subject, its verb must also be marked for plural. So although children appear to begin with inflections as modulations of the word meaning, they go on to learn the grammatical functions of each inflection, and hence the full domain for each one. Learning inflections ultimately demands attention to both lexical meaning and syntax.

When children learn inflections, they must also learn which words belong to which paradigms. That is, whether a noun takes the regular plural, or which two or three distinct regular plurals, versus an irregular form. In English, this typically amounts to a choice of a regular allomorph (/s/, /z/, or /ɪz/) versus an irregular form – for example, *-en* (as in *ox/oxen*) or a vowel change (as in *tooth/teeth*). Some languages have one highly regular paradigm and a scattering of small irregular ones, as in English plurals. Others may have several regular paradigms, with word membership in each dictated by phonological form, plus some smaller, irregular ones. In each language, children have to learn which paradigm a particular word belongs to for purposes, say, of plural formation. Where children have not yet learnt this, they may regularize some forms by assigning them to a major paradigm. At issue here is the extent to which they learn an inflection like the plural on a word-by-word

basis versus a constructional, rule-like basis. That is, once they have extracted the form of the plural affix, when can they add it, as needed, to any unfamiliar stem? Do children shift, at some point in development, from word-by-word acquisition of an inflection to rule-like application? And do all children go through a similar sequence of stages as they learn this?

Some affixes are more complex than others and are typically acquired later. Both meaning and form appear to contribute to their complexity. The complexity of the conceptual distinction underlying the meaning should affect how early and how readily children assign the pertinent conventional meaning to an inflection. In addition, the mapping of meaning onto an inflection may be affected by the complexity of the form itself. If the boundary between stem and affix is obscured by morphophonological rules, children will have a harder time identifying both, and so will acquire the inflection later. If they have to map a particular meaning onto discontinuous elements (e.g. a case ending and a preposition), that form too should take longer to acquire than one where the mapping involves a single element. In short, children must analyze the word forms they encounter, identify stems and potential affixes, and assign meanings to both. Their ability to do this may be taxed by the complexity of the meaning to be assigned, or by the complexity of the form for that meaning, or both. Only after this will they be in a position to produce the pertinent affixes.

Finally, are inflection and word formation treated in a similar way by children acquiring morphology? Inflectional paradigms tend to be complete, and so are generally more regular than word-formational ones. To what extent does this make word formation harder to acquire than inflection? Word formation also involves the construction of new forms to carry new meanings. Is constructing a new word harder than adding an inflection to a familiar or an unfamiliar one? In short, word formation seems to demand attention to, and hence knowledge of, the lexicon as a whole, unlike inflection. It also depends on knowledge of other typological properties of the language.

The discussion that follows will review some of the findings pertinent to these issues, and raise some additional questions about morphology in first-language acquisition. The next section focuses on inflections and their contribution to syntax. The third section takes up morphology in the lexicon through its role in new-word formation.

2 Inflections

Children's acquisition of inflectional morphology has been studied for a variety of different languages. Most of the data come from longitudinal records of children's speech. The data currently available come from a number of language families, including Indo-European (e.g. French, Spanish; English, German, Swedish; Polish), Semitic (Arabic, Hebrew), Uralic (Finnish, Hungarian), Altaic (Japanese, Korean, Turkish), Caucasian (Georgian), Australian (Walpiri),

Niger-Kordofanian (Sesotho), and Sino-Tibetan (Mandarin). For some of these languages, there are also systematic elicitation data available on some topics.

2.1 *What is acquired when?*

Children start to acquire inflections from their earliest word use on. But they may only produce them after some weeks, or even months. Consistent use of an inflection can be assessed against use in appropriate contexts on the one hand, and against use of other inflections (including use of no inflection) on the other. In highly inflected languages, children may produce their first contrasting inflections within two or three months of beginning to speak (e.g. in Hungarian, Turkish, Polish, Russian). But the production of the first inflections depends on several factors, including complexity and typology.

The typical sequence in the acquisition of an inflection such as the English plural suffix goes as follows: (a) no use at all in contexts that call for a plural form (hence *cat* in lieu of *cats*); (b) sporadic use on a few forms where a plural seems to be called for; followed by (c) general use and over-regularization (the inflection *-s* applied to *cat* and to words like *foot* and *man*); (d) identification of the relevant limits on use along with acquisition of irregular plural forms (*cat/cats* versus *foot/feet*). Some researchers have proposed that prior to learning the regular inflection for plural, for example, children learn the irregular forms, but then give them up in favor of regularized ones. But such a development seems quite unmotivated. If children have learnt the meaning of a form like *feet*, why drop it again to express the *same meaning* with a different form (*foots*)? An alternative account goes as follows. Children often identify irregular forms as base stems. For example, they may identify both *break* and *broke* as stems, or both *go* and *went*, without realizing, in either case, that these pairs each “belong” to just one verb. On this view, one prediction is that children will add regular inflections to both stems, and that is what they appear to do: they produce both *breaked* and *broked*, as if for two distinct verbs, and they produce both *goed* and *wented* (Clark 1987, Kuczaj 1977). What has not been investigated, though, are the meanings that children who inflect both forms might assign to a pair like *goed* and *wented*.

The order in which children acquire inflections has been studied in some detail for grammatical morphemes in English (R. Brown 1973, Cazden 1968). There, the best predictor of relative order is semantic complexity, with morphemes that are cumulatively more complex being acquired later. A morpheme marking *x* is acquired before one that marks *x + y*, and so on. This is consistent with Slobin’s (1973) identification of conceptual complexity as one major determinant of overall order of acquisition. What has not been established is a general conceptual base for measuring the complexity of specific morphological distinctions within or across languages.

A second major determinant of order of acquisition in production is formal complexity in the expression of a specific meaning. If a language marks *x* with

a single suffix, invariant across noun types, x should be simpler to acquire there than in a language where the same meaning is expressed through a combination of affix and preposition, with the shape of the affix also varying with the gender and number of the noun that it is attached to. Children should acquire the simpler type earlier than the more complex one, and they do. This can be seen in bilingual children's early expression of locative relations in Hungarian (early) versus Serbo-Croatian (late) (Mikeš 1967). In addition, children learn to produce a distinction marked by a regular inflection – where this applies to a large range of stems – earlier than the same distinction marked by a large number of different forms applying to small paradigms. This is the case, for instance, in children's acquisition of plural marking in English versus Egyptian Arabic (Omar 1973, Slobin 1973).

2.2 Rote learning, rules, and regularization

Do children need to hear each inflected form before they can use it? Or do they realize, after a time, that unfamiliar forms typically take the same inflections as familiar ones? That is, can they generalize about how to mark plurality, for instance, or past time, and then apply this knowledge to new forms? In 1958, Berko argued that if children learnt inflected forms by rote, they would be unable to add inflections to unfamiliar words. But, as she showed, five- to seven-year-olds readily added different inflectional endings to nonsense words. That is, children were applying a consistent procedure in marking a noun as plural or possessive, or a verb as ongoing or past. These procedures can be represented as rules for constructing the appropriate forms. (Whether such procedures involve templates or internalized procedures analogous to rules is unclear; see further Bybee and Slobin 1982, Clahsen and Rothweiler 1993, MacWhinney 1978, and Marcus et al. 1992.)

Once children have identified an appropriate affix to mark some distinction inflectionally, they can apply it wherever they wish to modulate a stem meaning in that way. They can add a past tense inflection, for instance, to any newly encountered verbs. The problem is that not all such verbs will be regular in form. So addition of the regular inflection will result in an over-regularized form such as English *breaked* or *doed* (for *broke* and *did*), or French *metté* 'put' or *pleuvé* 'rained' (for *mis* and *plu*). Children commonly over-regularize irregular forms during the earlier stages of acquisition. However, the extent, consistency, and contexts of such over-regularizations are in some dispute. Some researchers have observed them at very high rates (from 20 to 50 percent of the time at certain ages), but others, averaging rates across children and ages, have estimated them at no more than 3–10 percent (Kuczaj 1977, Maratsos 1993, Marcus et al. 1992). The issue is the following: do children simply go to a regularized form on those (perhaps rare) occasions when their memory fails, or do they go through a stage, during acquisition, where they assume that irregular forms are actually regular, and only later shift over to the conventional

irregular forms? If initial over-regularizations represent a stage in acquisition, they are liable to be more frequent initially for each irregular form and then to taper off as children begin to register that adults never produce the form that they themselves are using.

Which affixes do children identify as “regular”? Children latch on to some affixes very early, and others, with near-equivalent meanings, not until much later. Does frequency play a role here, and if so, do children attend more to token- or to type-frequency in the language they hear? Findings so far show that children attend more to types than to tokens in the input. The first affixes they produce are those that appear on the largest number of types (Guillaume 1927), and typically represent the most widespread paradigms in a language. Once these are established, children begin to acquire affixes marking smaller, irregular, and specially conditioned paradigms. But the latter can take many months or even years to master.

2.3 Case marking

In languages with case marking, children typically begin with just one form of each noun, generally the nominative or the accusative. Contrasting cases on the same noun in some languages begin to appear very early (around twelve to fourteen months), in others a few months later. One determinant appears to be the nature of the case system: where a single affix serves all forms of nouns, children master the case contrasts much faster, even with phonological conditioning, than where the forms of each case ending vary with the gender and number of nouns. Children show much earlier mastery of case marking in languages like Hungarian and Turkish than they do for German or Serbo-Croatian (Slobin (ed.) 1985, 1992).

The first contrast acquired seems to be between the nominative and accusative cases, associated with subject and direct object respectively. Contrasting uses of cases may appear with single-word utterances. In two-word combinations, where word order may offer no clues, case can distinguish the object of a transitive verb from the subject of an intransitive one. As children add other cases such as the dative and genitive, these too serve to distinguish direct objects, for example, from indirect objects (e.g. recipients and possessors) from the two-word stage on. In general, acquisition of nominative and accusative cases is followed by the remaining oblique cases. This may involve only two or three other cases in a language like German, versus many other case forms in one like Finnish.

Several factors make case difficult to learn. The most notable may be the number of forms that children have to deal with in some languages. For example, both gender and number interact with case. In a two- or three-gender language, there are typically multiple affixes for each case. And within each gender, languages may have several noun paradigms, with each paradigm identified by the phonological shape of the root or stem. Children have therefore to deal

with several different affixes as they learn how to express each case in a language. When it comes to number, they also have to deal with the fact that some gender distinctions in the singular forms are lost in the plural. In fact, children tend to learn first how to mark case in the singular, and only later in the plural. Languages differ, therefore, in the number of affix shapes to be learnt for each case, as well as in the number of cases – from a minimum of two or three to more than twenty. Where more shapes are associated with a particular case, children are more likely to opt initially for just one affix shape to mark a particular case on every stem. This reliance on a single affix shape, regardless of gender and number, has been dubbed “inflectional imperialism” (Slobin 1973). What this does is allow children to mark case with some consistency prior to the acquisition of gender or of subparadigms within genders. In languages where case affixes are invariable, or vary only, say, with vowel harmony, children acquire adult-like case marking very early, typically before the age of two – for example, in Turkish and Hungarian. In languages where case interacts with gender and number, children acquire the full system of case marking, with all the different affix shapes, much more slowly, and may still make some errors as late as age five or so – for example, in Russian. Adult-like case marking may also take more time in languages where complex morphophonological rules obscure stem–affix boundaries, and so make it harder for children to identify stems and affix shapes. Form, and in particular the range of forms for each affix, depending on gender and noun paradigms within genders, is a major determinant of how long children take to acquire adult-like case marking.

2.4 Person, number, and gender

Verbs are generally marked for person and number, and in some constructions and tenses for gender. The earliest verb forms used are typically third-person singular present, imperative, or infinitive in form. Children may focus on one or more of these as their earliest verb form(s), depending on the language being acquired. For example, in English, the first- or second-person present, the imperative, and the infinitive are all realized as an uninflected or zero-affix form of the stem, while the third-person singular present is marked by *-s*. Children begin with the uninflected form, and only later mark the third-person verb form in the present. In Portuguese, however, children favor the second- or third-person singular form as their starting point in verb use, and only later produce first-person forms (Simões and Stoel-Gammon 1979). Several factors probably contribute to children’s initial choice of a verb form: frequency in adult input and a tendency to make use of third-person forms in self-reference alongside some level of minimal inflection (compared to other verb forms). But the initial form favored by children differs somewhat with language, with choices converging on an imperative form alongside some

present-tense form, often in the third person (see languages surveyed in Slobin (ed.) 1985, 1992, 1997).

Number is marked in both verbs and nouns. In verbs, the plural forms are typically learnt some time after their singular counterparts. That is, children usually learn the singular forms for all three persons (first, second, and third) before they master the plural ones in languages that distinguish person and number in the verb. Number is mastered earlier in the noun than in the verb, and typically begins to be marked before age two. The distinction between one versus more than one may be signaled nonconventionally at first through modifiers such as *more* or a numeral (e.g. English *more book* for '(several) books', or *two magnet* for '(many) magnets'). Then children begin to add the regular plural affixes to nouns and to over-regularize irregular plurals. In English, for instance, children add the regular *-s* in lieu of the irregular forms required for nouns like *man*, *tooth*, *ox*, *child*, or *sheep*. In languages where there are a large number of irregular plural forms, children still identify the regular plural forms early and use them when they over-regularize (e.g. Clahsen et al. 1992). But irregular forms may take many years to master, with children continuing to make errors in their plural inflection as late as age twelve, as in Egyptian Arabic (Omar 1973). This shows that the formal complexity, in terms of the number of plural affixes and the conditions on their use, affects the point of acquisition for children. The smaller the number of affixes to be acquired in marking a distinction like plural number, the easier it is for children to master the adult options.

In adult speech, distinctions like plural are often marked redundantly in the sense that "plural" may appear not just on the pertinent noun but also on the accompanying demonstrative (e.g. *those trees*), and in the case of subject noun phrases on the verb (e.g. *Those trees have been cut down* versus *That tree has been cut down*). Three-year-olds acquiring English, for example, do better on a variety of comprehension tasks for the plural when they hear multiple or redundant marking of that distinction (Nicolaci-da-Costa and Harris 1983). This suggests that earlier uses of numerals or *more* may in part also reflect the forms used by adults for marking plurality that children hear in the input.

In mastering plurals, children often have to deal with gender as well, in nouns, and often in verbs too. Here again, factors related to the forms of gender marking appear to be an important determinant of how early and how easily children acquire gender marking. Where gender is marked consistently, with the same affix, for example, on the noun and on any adjective modifying that noun, children seem to find it easier to acquire. The same goes for gender marking in the plural: consistency in the form across nouns of the same gender, plus use of the same affix on adjectives and even verbs marked for that gender, makes for earlier acquisition, as in Hebrew (Levy 1983). But where form offers a less clear guide to gender marking, children take longer to master gender, and may rely initially on semantic rather than formal factors in adding the pertinent affixes, as in Icelandic (Mulford 1985). Similar considerations apply in the acquisition of Sesotho noun-class markers, which are also used

on any adjectives or verbs agreeing with the noun (Demuth 1988), and in the acquisition of noun classifiers in a language like Thai (Carpenter 1991).

2.5 *Tense and aspect*

Several researchers have suggested that children may initially assign an aspectual meaning of completion to past tense markers. One type of evidence is that children appear to be selective in which verbs they first mark with past tense inflections. In both English and Italian, they begin using the past tense on accomplishment (telic) verbs before other verb types (Antinucci and Miller 1976). One interpretation of these data is that children use the past inflection to mark the result or completion of the action referred to by the verb. This does not, of course, imply that aspect is easier to acquire than tense. What it does suggest is that the result or completion of certain action types is highly salient to young children. But in languages where both aspect and tense are marked on the verb, children appear to acquire both forms of inflection at an early age (starting well before two), with no clear ordering of the two (e.g. Weist et al. 1984).

The first tense contrast that children seem to introduce is that between present and nonpresent. The first nonpresent inflections usually mark completed, hence past, actions; but they may also mark future time. (In similar fashion, two-year-olds often use *yesterday* to mark either past or future (e.g. Decroly and Degand 1913, Harner 1975).) Slightly older children, around age four, commonly choose past-tense forms to mark *irrealis*. They do this in pretend play, for example, when assigning roles and planning future series of actions, as in the following exchange which preceded the relevant acting-out (Lodge 1979: 368):

A: I wanted to go.

B: But I wouldn't let you and you argued about it.

After acquisition of the initial present/nonpresent contrast, children add other tense inflections to mark the future and to distinguish past forms for background versus foreground events (typically, imperfect versus perfect forms). Some tenses such as the present perfect may not be fully mastered until age four to five, but the basic present/past/future contrasts are generally well established by around age three.

Aspect marking is acquired early in languages that mark aspect as well as tense through inflections on the verb. In Slavic languages, perfective and imperfective inflections, usually prefixes, appear at the same time as tense marking (Slobin (ed.) 1985). Aspect also appears to interact with the inherent aspectual meanings of verbs (Aktionsart). English-speaking children, for instance, add the limited duration suffix, *-ing*, initially only to activity verbs, and at first apply the past-tense suffix, *-ed*, only to verbs for change of state (e.g. Bloom et al. 1980; Clark 1996).

2.6 Agreement

One basic function of inflectional systems is to indicate which elements in an utterance “go together.” One finds agreement in number, person, and sometimes gender, for example, between a subject noun phrase and the verb, and agreement in number and gender between nouns and adjectives that modify them. There can also be agreement between articles or demonstratives and the nouns they go with – in gender, number, and case – and between pronouns (independent, possessive, or relative) and their antecedents – again in gender, number, and case. Agreement markers therefore help group together those elements that belong together for semantic and grammatical purposes. The acquisition of inflections must be measured, therefore, not just by the acquisition of specific paradigms, but also by children’s use of agreement more generally.

Overall, children appear to rely on phonological cues to gender and gender agreement. In French, for instance, children omit articles at the one-word stage, and they make some errors in their choices of article early on. But phonological form in French is correlated with gender (masculine or feminine), and children quickly become sensitive to such cues. In one elicitation study, where phonological form and natural gender were correlated, even the youngest children (aged three), having heard an indefinite article, produced the appropriate definite article nearly all the time. With phonological cues only, they did equally well; but with neither phonological clues in the shape of the word nor information about natural gender, they made errors in their choice of definite articles about 20 percent of the time. Where indefinite articles and word shapes conflicted (e.g. a feminine article with a masculine word shape), children up to six would change either the article or the noun shape to make the two agree (e.g. *le bicronne* would be changed to *la bicronne* or *le bicron*). Older children also took account of the natural gender of the dolls being labeled and assigned feminine articles and word shapes for female dolls (Karmiloff-Smith 1979).

Phonological cues to gender agreement also predominate in the acquisition of agreement in other languages. In Hebrew, for example, children rely on word shape in plural formation as they learn to apply *-im* (masculine) or *-ot* (feminine), and they extend these inflections to adjectives for noun–adjective agreement, without apparently taking any account of natural gender (e.g. Levy 1983). Similar observations hold for the acquisition of gender agreement in German, Polish, Russian, and Serbo-Croatian (see Ferguson and Slobin (eds) 1973; Slobin (ed.) 1985). In languages where there are apparently only minimal clues to gender in the actual word shapes, though, children instead seem to rely on semantic criteria, information about natural gender, in interpreting various forms of agreement (e.g. antecedent noun–pronoun). This suggests that children may have recourse to semantic factors in learning gender agreement only when the phonological clues are inconsistent (Mulford 1985).

Finally, in Sesotho, children again seem to focus on phonological cues to word type as they learn gender and agreement. As in other Bantu languages,

words in Sesotho belong to one of about fourteen word classes marked by prefixes that appear on the noun and corresponding forms for agreement on verbs, adjectives, demonstratives, and possessive, relative, and full pronouns. Children may begin to mark agreement with the noun on other elements, even before they can produce the noun-prefix reliably. This suggests that they are focusing on the whole noun or verb phrase in the input, so they treat the domain for agreement as some kind of prosodic and grammatical unit (Demuth 1988, Slobin (ed.) 1992). This would be consistent with children's trying to mark linguistic elements as going together if they help pick out the same entity or same activity together with its participants. That is, agreement is not simply an arbitrary set of markings: it has a readily accessible function for the language user.

2.7 Typology and acquisition

Children's patterns of acquisition suggest that they can process some kinds of information more readily than others. For example, they consistently learn suffixes before prefixes, even when these express equivalent information. Children learning a prefixing language like Mohawk acquire the inflectional prefixes later than children learning a suffixing language (e.g. Mithun 1989). This asymmetry is consistent with the more general asymmetry among languages: suffixing systems by far outnumber prefixing ones (Hawkins and Cutler 1988). When children are given nonsense prefixes and suffixes to imitate, they find suffixes easier than prefixes (Kuczaj 1979). In short, children seem to find it easier to process information added to the ends of words than to the beginnings.

Children also show a general preference for marking added meaning with an affix. For example, when presented with plural forms that differed from their singular counterparts (a) by the addition of an affix, (b) by subtraction of an affix, or (c) by zero, children prefer option (a), an added affix (Anisfeld and Tucker 1967). This is consistent with Greenberg's (1966) observation that added complexity (of meaning) is typically marked in languages by added morphemes.

Children learning different language types typically follow similar timelines, but several factors make for differences in the acquisition of inflectional morphology: (a) the consistency of the paradigms (Bybee 1991); (b) the nature of the meaning-to-form mapping, with one meaning to one affix shape being the easiest; and (c) the role of each inflection in syntactic constructions. Children exposed to a Semitic language first, for example, take for granted that it is the root consonants that provide the core meaning for each word, while those exposed to an Indo-European or Turkic language take both consonants and vowels into account in identifying words. That is, the effects of typology are relative, so they may be hard to assess. It appears possible, however, in bilinguals with languages of different types. Children growing up with Hungarian and Serbo-Croatian, for example, show that, to express the same

meaning, it is easier to use suffixes than a combination of suffixes and prepositions, and that consistency of meaning makes affixes easier to acquire.

3 Word formation

As children learn more words, storing them in memory and producing them themselves, they come to analyze their internal morphological structure. They begin to identify roots and stems inside complex words, in both compound and derived forms, and simultaneously isolate any derivational affixes attached to those roots. Such analysis is a prerequisite for new-word formation. And children do form new words, starting as young as age eighteen months to two years. In English, for example, they construct compounds and form verbs from nouns with no affixation. In the next few months (two and a half to three), they come to use affixes as well in the construction of new words.

Languages differ in the options they offer for coining new words. Some languages rely extensively, or even exclusively, on compounding; others rely mainly on derivation; and others rely on both. Are some options acquired more easily than others? If so, children learning different language types should follow different routes in their acquisition of word formation. The sections that follow review first what is known about children's acquisition of derivational options in word formation, and then their acquisition of compounding. Overall, children begin to use inflectional morphology before they coin new word forms, although there is considerable overlap in some languages. But derivational affixes in general begin to emerge later than inflectional ones.

Lastly, when children coin new words, they fill semantic gaps. Children do not wait until they have learnt the appropriate word before they try to express a particular meaning. Instead, when they need to, they construct a form for the meaning they want to convey. In doing this, they observe two general constraints on the coining of new words. Conventional words – forms that express meanings agreed on by the language community – take priority. If a word is already known to the child for the pertinent meaning, that is the word they use. And there is then no reason to coin another word with the selfsame meaning. New words must therefore contrast in meaning with existing words within any semantic domain (Clark and Clark 1979; Clark 1990, 1993). These two assumptions appear to be observed by both children and adults.

3.1 Derivation

Evidence that children are using derivation comes from their construction of novel words. To use an affix appropriately, for example, requires children to have analyzed that affix in established words and to have assigned it some meaning before they can use it in constructing new words. The first novel derived forms children construct are derived with no affix. Somewhat later,

around age three, they begin to produce an increasing number of novel forms with affixes.

Children coin new verbs in English from around age two. They form them mainly from nouns but also, on occasion, from adjectives, as in *to scale* 'weigh', *to key* 'insert a key', *to sand* 'grind', or *to water* 'paddle in water'. Such verbs require no affixation to indicate the change from noun to verb: they need only the appropriate suffixes and syntax (Bowerman 1974, Clark 1982, Maratsos et al. 1987). In effect, children are exploiting a zero-derivation option when they construct new verbs from familiar nouns. This effectively allows them to form new words from words with meanings already known to them, and to do so without having to make any changes in form. Even when they get older, children often omit to palatalize consonants, change vowels, or shift stress where these are required. For example, four-year-olds typically fail to recognize the root *magic* in *magician*, and when asked what they would call someone who does magic, the commonest response is *magic-man*. Equally, in forming an adjective from the noun *volcano*, five- and six-year-olds often construct *volcanic*, without the required change in the stressed vowel. Zero derivation is also favored early on in Hebrew, but there children use it to form new nouns from verbs (Clark and Berman 1984).

Between two and two and a half, children begin to produce derivational affixes, mostly suffixes, but these may be rare before age three. In English, for instance, children start to produce agent and instrument *-er*, though they do not use this suffix consistently until around age four (Clark and Hecht 1982). In other languages, the earliest suffixes to be used to form new words are generally diminutive endings and agent and instrument markers. Children typically show good comprehension of such suffixes for some time before they produce them themselves, and in elicitation tasks will produce other, non-derived forms instead. Finally, children also show distinct preferences for some derived forms over others with closely related meanings. The first agentive suffix they produce, for instance, is normally the one adults favor too (Clark 1993). For example, in English, children misremember novel agentive nouns as using *-er* even when they in fact have *-ist* or *-ian*, or even the non-agentive *-ly*, as a suffix (Clark and Cohen 1984).

These findings suggest that children rely on certain general principles as they analyze word forms and then construct new words themselves. They attend to the *transparency* of the components used; that is, they make use only of elements whose meanings they already know. This would account for why they initially rely on zero derivation, making use of stems or roots of familiar words. Only once they have assigned some meaning to an affix, do they begin to use that too in constructing new words. They also attend to the *simplicity* of the form produced; the fewer the changes to be made in the component elements, the easier it is to construct and produce. This again would lead children to favor zero derivation early on. And they are sensitive to the *productivity* of the affix being used; they follow adult usage in favoring the most productive option first, unless there is some reason not to (Clark 1993).

Finally, there is strong evidence that children analyze affixes and assign some meaning to them some time before they start to produce them themselves. For instance, when asked, they can offer glosses of what novel words might mean well before they will coin words using the same affix. English *-er* and Hebrew *-an*, for example, are readily identified as having agentive meaning by three-year-olds acquiring these languages, but not until age four or later do children produce those suffixes to mark agentive meaning (Clark and Berman 1984, Clark and Hecht 1982).

3.2 Compounding

In some languages, children begin to construct new compounds from as young as one and a half; in others, they make little use of novel compounds before age six or seven. The difference, in general, appears to depend on whether or not compounding is productive within the language. In addition, children are attentive to the *transparency* and *simplicity* of the elements they use in compounding. As a result, their earliest compounds typically consist of combinations of familiar bare nouns which are both transparent in meaning and simple in form (root compounds). It is only later that they begin to use affixes or produce any adjustments in form required for specific types of compounding in the language.

Children learning Germanic languages construct root compounds from an early age, often before age two. They form them mainly from familiar nouns, as in English *crow-bird* (one year, seven months, for 'crow'), *oil-spoon* (one year, eleven months, 'spoon for cod-liver oil'), or *coffee-churn* (two years, 'coffee-grinder'); German *Korb-wagen* (two years, seven months, 'basket-wagon', for a small doll's pram of woven straw), *Löchern-teller* (two years, nine months, 'hole(s)-plate', for a glass table mat with metal filigree on it), or Icelandic *kubba-bill* (two years, four months, 'block-car', for a car made of blocks), *flösk-bill* (same age, 'bottle-car', for a milk-truck), or *fiata-bill* (same age, 'Fiat-car'). And, by age two to two and a half, children have learnt to identify the modifier and head in such root compounds. In comprehension tasks, children acquiring English consistently identify the second (rightmost) noun as the head and the first as the modifier, while those acquiring Hebrew choose the first (leftmost) noun as the head and the second as the modifier (Berman and Clark 1989, Clark et al. 1985). But children acquiring languages that make less use of compounding do not produce compounds at this age. For example, children acquiring Romance languages produce virtually no root compounds until around age five or later (Clark 1993).

Children acquire synthetic compounds only rather later, and they often make errors in their construction. Elicitation tasks reveal that, in learning the structure of compounds like *button-thrower*, children seem to go through several stages. At around age three, they construct forms like *throw-man* (or, less frequently, *throw-button*) for the agent. They then begin to add the appropriate

affix to the verb stem, and now construct *thrower-button*. This form has the affix added to the head, and has incorporated the noun for the generic object affected; but the word order is the head + modifier characteristic of verb phrases (i.e. *throw* preceding *button*), instead of the modifier + head order of compounds. Only around age five do children acquiring English get both affix and order right when asked to coin such compounds. At that point, they begin to construct forms like *button-thrower* or *wagon-puller* with no errors (Clark et al. 1986). Children acquiring other Germanic languages make similar word-order errors, but children learning languages where the verb phrase head and compound head have the same order relative to any modifiers never make these word-order errors (Clark and Berman 1987).

Children rely on transparency and simplicity in their novel compounds just as much as in their derived words. Their earliest compounds are all root compounds, typically forms from two or more 'bare' nouns already known to them. They begin to form novel synthetic compounds only later, after analyzing and assigning some meaning to suffixes such as the English agentive *-er*. This sequence is predicted by both transparency of meaning and simplicity of form. Further evidence for simplicity comes from the sequence observable in the acquisition of Hebrew compounds. In Hebrew, different noun types differ in their bound forms, the ones required in compounds: some make no change; feminine nouns in final *-a* add a *-t*; masculine plural nouns change final *-im* to *-ey*; and so on. In production, three-year-olds favor no change in their compounds, and so make many errors. As children get older, they successively master the final *-t*, then plural *-im* to *-ey*, and, last of all, stem-changing heads. That is, the greater the change required in the form of the head, the longer before it is acquired (Clark and Berman 1987).

Productivity also plays a role in compounding. Children favor the patterns that are more productive over those that are less productive, or no longer productive. This suggests that here too children attend to the relative frequencies of different compound types in the input around them. In level-ordering models of word formation, all compounding is assumed to occur at the same level with all of it productive. But contemporary speakers favor only certain patterns among those possible, and the most productive of those are the ones that children typically acquire first. Frequency, then, presumably makes for accessibility during acquisition.

4 Conclusion

In general, children start to acquire inflections before they begin on novel-word formation. The earliest noun and verb inflections to emerge appear in some languages before age one and a half. Compounding with no affixation emerges soon after the first inflections, but novel derived forms do not emerge until after age two. The first to appear are zero-derived forms with no affix.

Then come some derivational affixes, with sporadic use up to age three, followed by more extensive use from age three or four on. In languages that make little use of compounding or zero derivation, therefore, the first novel-word formations may not appear until age three or later. It is unclear whether typology affects the acquisition of morphology elsewhere. Although it appears easier to process suffixes than prefixes, there are too few data on the acquisition of prefixing languages to see how consistently this holds overall. At the same time, children acquire locative affixes, for example, much earlier in languages that use invariant forms on all stems than in languages that rely on a mix of case marking (varying in form with gender and number) and prepositions. It may be easier in general, then, for children to map inflectional meanings in agglutinative than in synthetic languages.

Overall, the sequence of acquisition for morphology, whether in inflectional systems or in word formation, appears to depend on at least two factors: the complexity of the meaning being expressed – where children have to discover this for each affix – and the complexity of the form to be used – where children have to work out the conditions that govern different allomorphs. However, what counts as easy versus difficult in adjusting the form of a word is not easily measured.

In short, children work with words. Their earliest inflections are typically learnt as parts of words, and only later are analyzed for forms and meanings. Once this is done, children appear able to extend paradigms with rule-like application of an affix to new instances. In doing so, they also regularize irregular forms until they learn to produce the appropriate irregular forms. This holds for both inflections and word formation.