12

The acquisition of grammatical categories

Heike Behrens

12.1 Introduction

Grammatical categories encode conceptual domains and contrast them with others. They include the part-of-speech categories of words as well as morphological markers that encode functions like number, case, gender, tense, aspect or voice. In linguistic terminology, morphology deals with the grammar of words (Booij 2012), whereas syntax is concerned with the relationship between words. Words differ regarding the syntactic functions they can assume (part-of speech category or word class), as well as their morphological structure: a word stem can be combined with inflectional and derivational affixes, or undergo stem changes to encode grammatical functions; and words can also take part in compositional processes to form new, more complex words (compounding, derivation). Grammatical morphology is typically encoded by affixes that combine with content words like verbs, nouns and adjectives.

Typological research shows that languages differ widely regarding the categories they differentiate by means of morphology, and the way in which they do it (see Chapter 6). The language-learning child has to segment complex morphological forms in order to identify the stem, but also to identify the function of the grammatical markers, and establish inflectional paradigms. When the system is fully mastered, the speaker is able to use the lexical items as well as the inflectional markers in the appropriate range of constructions, while also observing constraints in the use of these forms.

Section 12.2 provides an introduction to the formal and functional specification of grammatical categories, followed by a characterization of the learning strategies needed to acquire complex grammatical forms (Section 12.3). After a discussion of typical morphological development and the criteria to assess development (Section 12.4), processes in acquiring the

form–function mapping in plural formation and tense–aspect marking will be discussed (Section 12.5).

12.2   What are grammatical categories?

When grammatical categories are conceived of as linguistic encodings of certain concepts or functions, both the nature of the concepts or functions and the nature of their formal encoding need to be investigated. In this section, a short characterization of crosslinguistic research on grammatical categories is provided.

12.2.1   Function

Languages typically have open-class, often free-standing elements like verbs and nouns, and closed-class elements like function words (e.g. prepositions, topic markers) and morphological affixes that encode tense, aspect, case and numerous other distinctions. Grammatical relations are typically encoded by word order and/or closed-class elements, and hence typologists analyse whether there is a principled distinction regarding which aspects of language are encoded by grammatical categories. This question is linked to the issue of whether some aspects of the grammatical structure of human languages are innate, such that these notions are encoded first or in a very stable manner in early child language.

Closed-class items include, in particular, tense, aspect, causativity, voice, mood and person, whereas other notions do not seem to be encoded grammatically, although they are part of our everyday experience: for example, languages do not tend to encode colour grammatically, nor the state-of-mind of the speaker, i.e. whether he was interested or bored (Talmy 1985; see also Slobin 1997b). In addition, grammaticalization processes in different languages seem to show similarities which suggest that the processes and representations are common to all humans and not the product of an individual mind or culture. If there was a systematic conceptual divide between lexicon and grammar, two different acquisition mechanisms would be needed: one for identifying the – potentially unlimited – semantic richness of open-class lexical items, as well as a look-up mechanism for the semantically constrained set of closed-class grammatical categories (Slobin 1997a: 267). Subsequent typological research on crosslinguistic differences has shown, however, that form and meaning are not neatly carved up into entirely predictable packages of closed-class items with a constrained set of meanings, and lexical items with unconstrained semantics (see, for example, the online *World Atlas of Linguistic Structures* for the distribution and encoding of a range of grammatical distinctions crosslinguistically: http://wals.info; see also Chapter 6).

While languages show considerable variation in how they encode certain notions, it is possible that children across the world start learning their language(s) based on similar functions, if not same semantics, because, for example, (a) they perceive visual stimuli in the same way in spatial configurations (Spelke & Kinzler 2007); or because (b) they are involved in similar events and activities, for example, scenes where an agent manipulates an object (e.g. Slobin 1985a); or because (c) their general cognitive development proceeds in a similar fashion. For example, it should be cognitively easier to refer to past events that resulted in an observable change-of-state than to encode temporal relations like before and after (Li & Shirai 2000), which may involve a different deictic centre from the here-and-now, and the order of events may be different from the order specified in the utterance (e.g. *Before John went to Paris he spent five days in London*; Clark 1971).

Crosslinguistic research has shown that not only do languages differ in how they divide up 'semantic space', but that language itself plays an important role in shaping children's linguistic and nonlinguistic categorization (Bowerman 1985, 1994, Gentner & Bowerman 2009, Slobin 1997b; see also Chapter 6). Children need to pay attention to those categories that are (obligatorily) encoded in their language: Turkish children, for example, need to keep track of whether an event they report on has been observed directly or not, since Turkish encodes evidential modality (Aksu-Koç 1998). From early on children reproduce the form–function relations found in their target language, rather going through a phase shaped by general, language-independent notions (e.g. Bowerman, Brown, Eisenbeiss, Narasimhan & Slobin 2011, Choi, Bowerman & Mandler 1999; but see Gleitman & Papafragou 2012, and Göksun, Hirsh-Pasek & Golinkoff 2010 for similarities between children in the discrimination of spatial concepts in the prelinguistic phase).

12.2.2   Form: crosslinguistic differences and 'typological bootstrapping'

The term 'grammatical' refers to the syntactic and morphological relations in a language. The distinction between syntax and morphology is not an absolute one, because some languages encode morphologically what others encode by specific syntactic constructions or lexically through circumscription. For example, English forms the passive by a syntactic operation as well as morphology (*Peter hugs his dog → The dog is hugged by Peter*), Latin or Sesotho do so by inflection. Thus, languages differ in what is encoded by syntactic means and what is encoded by morphological means.

Morphology and syntax are related to phonology as well as semantics and pragmatics. Slobin (1997b: 277) defines a grammatical morpheme as follows: prototypical grammatical morphemes are affixed to content words, are general in meaning, phonologically reduced, and not etymologically

transparent. Familiar examples are elements like plural markers on nouns and tense/aspect inflections on verbs. Phonological and prosodic properties of the stem may determine the declension class of a word or the allomorphs. For instance, the English plural morpheme has three different realizations (*-s, -z,* or *-əz*) depending on the noun stem (e.g. *cat-s, pig-s* and *fox-es*). Likewise, grammatical categories can encode semantic distinctions, e.g. tense and aspect, or pragmatic ones, e.g. honorifics (different morphemes to encode degrees of politeness).

From a typological perspective, languages differ widely in their division of labour between syntax and morphology, and in the number of categories they distinguish morphologically. Fixed-word-order and isolating languages (e.g. Chinese) have no grammatical morphology and few function words, whereas richly inflecting languages (e.g. Russian) can have highly complex morphology where words have several markers. Turkish represents the agglutinative language type where affixes that encode grammatical categories are added to the stem. Each grammatical category is typically represented by a different morpheme. In Turkish, we find a very regular type with a one-to-one mapping of form and grammatical function. Walpiri, in contrast, has allomorphs for certain categories like the ergative. But inflection can also be fusional or synthetic. In fusional languages, inflectional morphemes typically encode several functions (e.g. case plus number or tense plus agreement in German), and the form of the affixes can vary depending on the declension class of the stem, as in Russian. Finally, in (poly)synthetic or incorporating languages not only grammatical affixes can be added, but also other words or their pronominalized forms. Thus, a number of elements are incorporated into a single word. Incorporating languages include Inuit languages, indigenous languages spoken in the Americas like Mayan, or Sino-Tibetan-languages like Chintang (see Chapter 13). This variation implies that the number of obligatory or optional grammatical categories possible in a language as well as the type of the morphology (additive or fusional/synthetic) will affect what the child will have to pay attention to when learning to segment and process inflected units. For inflectional morphology, children will need to notice the contrasts between different inflected forms of a verb (or noun), and establish a form–function mapping in making a first analysis of what each contrast may encode.

Slobin's (1973) study of cognitive prerequisites of language established systematic crosslinguistic comparisons of language acquisition. He operationalized differences between languages in order to find out what makes the acquisition of a language easy or hard. He found, for example, that children learning Turkish, a very regular agglutinative language, acquired certain grammatical notions much earlier than children learning Serbo-Croatian, a language with highly complex inflectional morphology (see also the chapters in Bavin & Stoll 2013, Berman & Slobin 1994 and Slobin 1985–97 for systematic comparisons of the acquisition of morphology and

syntax in typologically different languages). In order to understand more about these generalization processes, two questions have to be addressed: how can we measure children's morphological development, and how can we assess the productivity of their linguistic representations, that is generalisation of grammatical forms?

12.3 Generalization and schema formation

In this section we focus on the onset of morphological paradigm building. In a usage-based perspective, generalization over experience leads to schema formation (Langacker 2000). Thus, generalization is a bottom-up process. On the one hand, repeated encounter of particular strings in a language leads to entrenchment in that the memory traces for that string become more robust. With an increasing amount of stored linguistic experience, the child will start to register commonalities between certain forms and the functions they encode. For example, the child may start to register that words denoting events or activities receive the suffix *-ing* when referring to ongoing events or activities, but the suffix *-ed* when referring to the past. The child can now relate the different forms of one stem to one another, but can also relate all forms marked by *-ed* or *-ing*, thus forming grammatical categories. Acquisition then is a process from low-level schemas where the related strings still share a lot of surface similarity, to more abstract schemas.

From the perspective of a child engaged in communicative acts, the child is usually confronted with whole utterances, not individual words or morphemes. Tomasello (2003: 174) identified the emergence of grammatical categories as a process of identifying their function in larger communicative units. In doing so, they typically move from concrete item-based, lexically restricted constructions to abstract representations which allow them to apply their knowledge productively in new contexts:

1. On the level of an expression, children have to identify communicative intentions in the input (e.g. *I wanna see it*), and they have to be able to reproduce these expressions.
2. Children form a pivot schema by forming a schema on the one hand, and a slot-filler category on the other (e.g. *throw ball, throw can, throw pillow* where the verb *throw* has an open object slot).
3. Next, children form item-based constructions as second-order symbols (e.g. *Mary hugs John, John hugs Mary*).
4. Then, children form abstract representations like syntactic roles or grammatical categories based on analogy (*A hugs B, X kisses Y*).
5. And they form paradigmatic categories (e.g. part-of-speech categories or inflectional paradigms) based on categorization by distributional analysis.

Bybee (1995, 2010) attributes this pattern of development to the different roles of type and token frequency: while high token frequency leads to strong entrenchment, high type frequency leads to generalization: e.g. children will note that the vast majority of 'verbs' in English form the past tense by adding *-ed*, and thus segment *-ed* as a morpheme and apply it to new verbs. This also helps explain why high-frequency irregular forms (like the English irregular verbs) remain in the language although their inflectional pattern is no longer productive, and why they may be acquired early despite their irregular status (see the results from morpheme-order studies, Section 12.4.3). Experimental evidence shows that children are influenced by type frequency: Dąbrowska (2006) tested this in diminutive formation in Polish children and adults by asking them to produce diminutives with nonce nouns. The 2-year-old children were better with those nonce nouns that fit into a dense phonological neighbourhood, where there are many nouns that show the same phonological patterns, and were somewhat worse with producing diminutives on words that had fewer similar neighbours. Adults did not show this effect, which suggests that young children's production of diminutives is not yet fully abstract.

Apart from the frequency of lexical items and affixes, recent studies suggest that children also rely on string frequencies, not just on the frequency of word forms (see Bannard & Lieven 2012). In a sentence completion task, Arnon and Clark (2011) demonstrated that children were better in producing irregular plurals in English when they had to complete a lexically specific frame (*brush your* teeth, *on your* feet) then when they had to label a picture (*What are all these?*). Likewise, distributional cues help children to induce part-of-speech categories: from corpora of child-directed speech, Redington, Chater and Finch (1998) found analyses of trigrams (strings of three words) lead to clusters of words that belong to the same category. For example, words that occur between *he* and *the* are typically verbs (*He ___ the*; see also Elman (2004) for more fine-grained analyses that lead to semantically based clusters). However, the reliability of distributional cues is language dependent: whereas Mintz (2003) showed similar effects for English and French, Erkelens (2009) and Stumper, Bannard, Lieven and Tomasello (2011) found trigram information alone somewhat less accurate for grammatical category induction in Dutch and German, where word order is more variable. This suggests that, depending on the native language structure, the child may have to focus on different aspects of the distributional information. However, in all languages studied here, a Zipfian distribution was attested where relatively few frames constitute a large portion of the input corpus. This skewed distribution contributes to learnability, because informative information is frequent and thus readily available and easy to compute. These analyses of the distribution and frequency of linguistic strings inform us about item-based, string-based and frame-based information that is

available in the input, and which forms the basis for children's generalization from low-level to more abstract generalizations (see Section 12.6).

The following section discusses studies on the emergence of morphology and on measures of productivity.

12.4   The acquisition of inflectional morphology

12.4.1   Assessment of development

Acquiring morphology can be an additive or a deconstructivist process. In the additive scenario, the child starts out with an uninflected stem form and learns to add morphological markers, and builds up the inflectional paradigm for each grammatical category. This order of acquisition is typical for languages like English with relatively little morphology, and where the stem of a word corresponds to the citation form, e.g. *laugh*. The child then adds agreement, aspect and tense morphology, e.g. *laugh-s, laugh-ing, laugh-ed*.

In the deconstructivist process, children start out with complex morphological forms and analyse their internal constituency only later. They have stored the form holistically, just like a second-language learner who may know a greeting formula in the new language without being aware of its structure. This is typical for richly inflected languages where children never or only rarely encounter stems in isolation (Xanthos *et al.* 2011). But it is also found in other languages, where children appear to use inflected forms or even larger chunks in an unanalysed fashion. Consider the closest relatives of English: German and Dutch. Here, the infinitive (the citation form of the verb) has an inflectional suffix *-en* (*mach-en* or *mak-en* 'to make'). Strictly speaking the form is compositional, which raises the question of how to conceive of such forms when children first produce them, if there is no evidence that the child is aware of their internal structures: to assume at face value that the child has access to the compositional structure of a complex morphological form would be an overestimation. Thus, criteria for productivity are needed; these are discussed in Section 12.4.2.

The assessment of productivity also affects a more general measure of linguistic complexity, the mean length of utterance (MLU; Brown 1973). For the early stages of language development, the MLU has been identified as a more reliable measure of children's language development than age alone. As a measure of growing linguistic complexity, MLU works best in languages where acquisition is predominantly an additive process. In English, the early stages of inflectional development are characterized by adding morphemes to stems (e.g. *Daddy walk* → *Daddy walk-s*). Brown computed MLU in morphemes, and later studies confirmed that in English, there is a high correlation between MLU measured in morphemes and words (Parker & Brorson 2005). However, measuring the MLU in morphemes can result in potentially inflated MLUs for children learning

languages with a richer inflectional repertoire (e.g. Rivero & Golędzinowska 2002). Here, children rarely encounter uninflected forms, and reproducing inflected forms that they have not yet segmented and analysed would result in an inflated MLU. Thus, MLU is now more commonly measured in words; however, this in turn renders a potential underestimation problem for learners of polysynthetic languages whose 'words' consist of many morphemes and have a sentence-like status (Fortescue 1984), or for synthetic languages and languages with portmanteau morphemes, where one stem change or affix can carry several functions (Dromi & Berman 1982; see also Chapter 13). Today, several measures of utterance length and internal constituency are usually analysed, depending on the research question. For example, one can measure utterance length as length of vocalization, or length in syllables, in addition to the traditional measures of words or morphemes (see Allen & Dench in press, Lyon & Gallaway 1990). While these measures can provide us with a general developmental index regarding the child's increasing utterance length, they do not provide us with a reliable measure of the degree of analysis or productivity of their utterances, as do the measures that are discussed next.

12.4.2   Criteria for productivity

A major problem in the study of child language is to determine when an item has been acquired by the child. The most straightforward criterion for acquisition seems to be the adult-like provision of the morpheme(s) under investigation. In his seminal study on the development of language in three American children, Brown (1973) set 90 per cent provision of a grammatical morpheme in obligatory contexts as the criterion for acquisition. This criterion was met when the child provided a morpheme (e.g. past tense or agreement) in 90 per cent of the utterances where it was required. The problem here is to define obligatory context: the linguistic and noninguistic context has to be taken into account in order to decide whether, for example, the use of a particular tense marker or the definite or indefinite article is required (Brown 1973: 255–256 and 259–270, Cazden 1968). The 90 per cent criterion of provision in obligatory contexts measures the endstate of development.

Before that, children typically show limited productivity in their language, either because they fail to provide certain grammatical markers when required, or by using them in a limited range of contexts only. In order to assess children's growing productivity with language, experimental techniques (see Ambridge 2012, Ambridge & Rowland 2013, McKercher & Jaswal 2012), comprehension data (see Chapters 8 and 13) as well as production data can be used. In language production, lack of productivity manifests itself in frozen or formulaic usage: The child might use complex forms correctly, but without evidence for having segmented the form and having started on category formation. For example, children could use the plural

form *car-s*, but unless the child contrasts this form with the singular *car* or uses the plural *-s* in other contexts as well, we cannot be sure that they have started to note that the *-s* affix encodes plural. In the context of child language research, productivity refers to the child's internal analysis of a form, not the productivity of the morphemes as such, i.e. whether they are synchronically used to inflect new words that enter the language. In child language research, productivity measures analyse whether children are aware of the internal structure of a string of their language, and whether they have developed generalizations about the internal structure of the string in order to apply them to new contexts where they cannot rely on a rote-learned form. For example, when children produce overgeneralizations like *run-ed* instead of *ran*, this is a very powerful diagnostic for knowing the function of the inflectional marker (see below for related evidence in experimental design with nonce words).

At the onset of development, the criterion of contrast or alternation marks the beginning of paradigm building (Section 12.3.4; see also Chapter 6). Here, the child has to use at least two different inflectional forms with the same word stem as the first sign of productivity. For example, a child uses *car* as well as *car-s* in order to mark the distinction between singular and plural, or *walk* vs *walk-ed* to encode a tense distinction. Alternatively, one could look for the occurrence of a morpheme with different words, like the plural morpheme used on two different nouns. The criterion of contrast between two forms is not a very strict one. Hence, some researchers prefer a criterion of at least three to up to eight (Braine 1976) different forms in order to reduce the chance that the child has simply memorized two or three inflected forms without having analysed their morphological constituency.

However, setting a threshold of a fixed number of contrasting forms is somewhat arbitrary. The main reason is the indeterminacy of the notion of 'onset of productivity'. While having a contrast of three or four or five forms seems better than having a contrast of just two forms, there is no absolute dividing line between acquired/non-acquired. Thus, setting a number is just an operational criterion to fix the minimum number of exemplars. A second reason is that a low number is often used because of the nature of the dataset, in particular its sample size (Rowland & Fletcher 2006). In many cases, it is just not possible to find more than two or three relevant examples within a single transcript, either because the inflectional category is rather rare (e.g. the passive) or because the inflectional category shows a lot of allomorphy such that the individual allomorphs have low type and token frequency. For example, in many spoken dialects of German, the past tense has been replaced by the present perfect, and is used only in written genres or narratives. Depending on the situations being sampled, children's command of the past tense could be underrepresented. If the productivity of a certain form is at issue, dense sampling (Lieven & Behrens 2012), targeted diary studies (Naigles 2012,

Naigles, Hoff & Vear 2009), carefully designed games (Eisenbeiss 2009), or longitudinal elicitation studies (Rowland & Theakston 2009, Theakston & Rowland 2009) can be used to obtain a large enough dataset.

Yet other criteria for acquisition could be set by comparing the frequency or the semantic and pragmatic appropriateness of the child's use of a morpheme with that of the adults (Brown 1973: 255). It is often observed that children use certain markers correctly and even contrastively, but omit them more frequently than adults would. English-speaking children, for example, omit subjects more frequently than adults, and the pattern of omission gives rise to theoretical debates (Graf, Theakston, Lieven & Tomasello 2015, Valian, Solt & Stewart 2009). In the case of tense and aspect marking, children's use is semantically constrained, and they acquire the fine-grained semantics of tense marking gradually (see Section 12.5.2).

More clear-cut examples for productivity are children's errors, most notably in the form of overgeneralization, when an incorrect allomorph is applied, e.g. a regular form for the irregular one (e.g. *go-ed* instead of *went*). Since it is unlikely that the child hears *goed* in the input, she or he must have made up the form. Such effects can be reproduced experimentally by testing children on nonce words, for example in the famous *wug* tests by Berko Gleason (1958). She presented children with unfamiliar nouns and verbs and elicited plural, past tense or progressive forms. The use of nonce words ensures that children cannot rely on their memory because they have not heard that word before. To pass such tests, children have to form a representation of the nonce word in order to produce it with the appropriate grammatical morphology, i.e. the task demand is higher than when inflecting a familiar word. Thus, nonce words tests estimate the upper end of productivity rather than its onset.

In their study on early passives in Inuktitut, Allen and Crago (1996: 139–43) discussed several less frequently used criteria to assess productivity. Innovative forms created by conversion of a word to a different word class reveal productivity; for example, as in *It ball-s* in response to the question *What does the ball do?* Likewise, self-corrections can also serve as an indicator that the child is aware of morphological forms. In languages with several affixes, (errors in) the ordering of the morphemes can also serve as a diagnostic. This would be the case if, for example, children were aware of the scope effect of the different positions of the passive morpheme in an agglutinating language like Inuktitut (Allen & Crago 1996: 141).

It should be clear from this discussion that the assessment of productivity involves a thorough analysis of the linguistic domain under investigation, as well as a thorough investigation of the developmental processes involved. This also involves theoretical considerations regarding the nature of children's linguistic competence (see Ambridge & Lieven 2011 for a discussion of the acquisition of several morphosyntactic domains under different

theoretical accounts of language acquisition, and how their predictions can be tested by comparing acquisition in languages that show a different alignment of the relevant form–function pairings).

12.4.3   Morpheme order studies

The order of acquisition of grammatical morphemes has been a major topic in acquisition research. Within each language, morpheme-order studies define the typical pathway of acquisition in terms of the time frames in which (clusters of) morphemes emerge (see Brown 1973 as well as the articles in Slobin 1985–97). Morpheme-order studies are largely descriptive, and do not offer direct insights into acquisition processes on their own. From a crosslinguistic perspective, however, differences in the order of acquisition of 'similar' morphemes can provide information on cognitive, phonological, semantic and distributional factors that influence children's segmentation and classification processes. For example, grammatical categories may or may not be obligatory in a language, can be more or less regular, and may show different types of selectional restrictions in terms of their distribution. Such variation can lead to different courses of acquisition for related concepts (see Section 12.5.2).

12.5   Form–function mapping in learning grammatical categories

Grammatical categories can pose very different acquisition problems. In some cases, the function that is encoded is relatively simple and uniform, but their encoding can be more or less regular. Such is the case with plural marking. The semantic distinction between 'one' and 'many' is known to children even before they start to learn inflection (Clark & Nikitina 2009), but languages differ in the degree of regularity of the inflectional paradigm. In other cases, the function itself can be complex, and be encoded by different markers. This is often the case in tense and aspect marking. Here, languages differ considerably in the number and nature of distinctions that are encoded, and how the grammatical categories interact with one another, as well as with the lexical properties of the verbs they combine with. In addition, different forms of marking can interact to encode certain grammatical relations. Notions like transitivity or transfer can be encoded by word order and/or case marking, for example. Such domains let us study which combinations of factors can be learned more easily.

12.5.1   Plural marking

Number is typical grammatical category marked on nouns. Apart from a general distinction between singular and plural, some languages like

Walpiri or Hebrew also distinguish the Dual to encode units of two, typically entities that appear as pairs. For languages that distinguish singular and plural, plural marking is conceptually easy because the difference between one or more items is visible and often communicatively relevant. Thus, children are likely to encode plural early, either by means of periphrasis (*more car, another one car*; Clark & Nikitina 2009, Zapf & Smith 2008), or by early generalization of grammatical plural markers (Zapf & Smith 2007). Plural marking on the noun is marked by inflectional affixes (*two duck-s*) or stem changes (e.g. German vowel raising in *Apfel – Äpfel* 'apple – apples'). When a language has plural affixes, there are often different variants or allomorphs: which allomorph is chosen depends on the phonotactic properties of the stem. In English, the plural morpheme is -*s*, and its acoustic realization depends on coarticulation: it is voiced after voiced consonants or vowels (*dog-s* = /z/), voiceless after voiceless consonants (*cat-s* = /s/, and syllabic after nouns that end in a sibilant (*fox-es* = /iz/). But plural marking can also interact with other categories like gender: in Hebrew, there are different affixes for masculine or feminine nouns. In addition, affixation leads to stress shift to the suffix (e.g. *tik – tikím* 'bag-s'). If the stem ends on a vowel, the affix replaces the vowel (*sirá – sir-ót* 'boat-s'; for further details for other noun types see Ravid *et al.* 2008: 37 f.). Thus, children have to work out whether the language has plural morphology, and which phonological, metrical and grammatical properties determine the choice of allomorphs.

A prominent proposal in the acquisition and processing literature is the Dual Route Model, which proposes two neural mechanisms for the processing of morphology, a rule-governed one and a memory-based one (Clahsen 1999a, Marcus 2000, Marcus *et al.* 1992, Pinker 1999, Pinker & Prince 1991). Irregular forms would have to be stored in memory, whereas regular forms could be generated from the stem by a regular rule. For example, English has very few irregular plurals (e.g. *child-ren, ox-en, fish-0*) which would have to be stored by rule. All regularly inflected forms could be generated by a symbolic rule that adds an -*s* affix to the noun stem. In actual language processing, the speaker would search through the stored inventory of irregular forms, and if the item is not found, compose the regular form by rule. Failure of look-up results in overgeneralization errors like *fish-es*. The Dual Route Model reduces the need for storage in the mental lexicon because all regular forms can be generated by the respective rules. The original version of the Dual Mechanism Model was much criticized because it did not account for morphological systems where instead of marking a clear division between regular and irregular forms, the inflectional system is more complex and conditioned by several factors (see below). More recent versions make a distinction between a declarative, memory-based processing system and a procedural one for the implicit learning of new items as well as the control of established cognitive processes like rules

and schemas, but no longer links these systems to regular vs irregular morphology (Ullman 2013).

In contrast, Single Route Models do not make a categorical distinction between regular and irregular morphology, but assume that all degrees of regularity can be learned on the basis of generalizing over the input children hear (Elman *et al.* 1996; see also the discussion in Clahsen 1999). The German plural system provides an interesting test case in this debate: German has four different plural suffixes *(-s, -er, -(e)n* and *-e)* as well as zero marking of the plural, and combinations of zero-marking or affixes with vowel raising (Behrens 2002, Szagun 2001). These allomorphs are more or less predictable from the properties of the noun stem: for example, all feminine nouns that end in *-e* take the *-n* plural. Nouns that end in full vowels tend to take the *-s* plural (*Auto-s* 'car-s'). The first is a high-frequency class, the latter a low-frequency class with mostly borrowed or truncated words. Although these classes differ in frequency, the plural forms are all highly predictable, and errors are rare (Behrens 2002, 2011, Szagun 2001, 2013). Other classes are less predictable: monosyllabic masculine or neuter nouns take the *–e*, *-er* or *-s* plural, sometimes with and sometimes without vowel raising. Here, the error rates are higher, but the errors are typically within the range of possible plurals for this type of noun (Behrens 2002, Szagun 2001, 2013). Behrens (2002, 2011) investigated the error rates and error patterns in a German child's learning of the plural in German. Throughout the first six months of morphological development, some subsets of nouns were not susceptible to errors. For example, nouns ending with a schwa always take the *-n* plural, and the child never made an error on this class. Other, less predictable, plural classes showed much higher error rates, irrespective of type frequency of that plural class, probably because the child at this early age did not have enough linguistic experience to generalize based on type frequency, and because some of the more uncommon allomorphs apply to nouns with high token frequency. Crucially, the error patterns changed with development: initially, the child overgeneralized in particular the *-n* ending (e.g. *Füss-e-n* instead of *Füss-e* 'feet'); a little later, he made frequent errors with nouns that do not take a plural suffix, because these nouns violate what is referred to as the iconicity principle – that plurals should be marked (Köpcke 1998).

In addition to the properties of the source (i.e. the phonotactics of the root), children also draw on the Gestalt or schema of the product, i.e. the inflected form. In a product-oriented generalization, the child would adjust forms to the schema of the inflected forms in the adult language. For example, in German a string consisting of the article *die* 'the' followed by a bisyllabic noun ending of *-(e)n* has a very high likelihood of encoding a plural (*die Frau-en* 'the women'), whereas the articles *der* and *das* only occur in the singular, such that these articles together with a monosyllabic noun would be a typical singular schema (e.g. *das Boot* 'the boot). When children make errors in their plural formation, the resulting forms typically

correspond to good plural schemas. Children might try to match their own plural forms to the schema 'bisyllabic noun ending in -*(e)n*'.

This product or schema orientation explains certain truncation of nouns that cannot be accounted for by a simple addition of inflectional morphemes (Behrens 2002, Bittner 2000, Köpcke 1998). For example, the German words *Apfel* 'apple' and *Vogel* 'bird' form the plural by vowel raising of the stem, without adding a suffix (*Äpfel, Vögel*). This is an ambiguous German plural from the schema perspective, because the ending is ambiguous between singular and plural forms. Children produce novel plural forms by adding suffixes (**Apfel-n, *Vogel-n*) but also by truncating the stem (**Apf-en, *Vog-en*). In the case of truncations, it is clear that children do not just add suffixes to a stem, since the stems are *Apfel-*, not *Apf-*, and *Vogel-*, not *Vog-*, but they cut of the final consonant to produce a plural that corresponds to the clearest plural schema in German. Likewise, they may double mark the plural by adding an -*n* to the correct plural affix -*e* because -*n* forms are 'better' markers (e.g. **Fisch-e-n* 'fish', **Füss-e-n* 'feet'). Overall, German children identify the plural allomorphs within a few months after their first plural use, as shown by early productive overgeneralization of all plural allomorphs (Behrens 2002). But it takes them several years to fully acquire the system, because they make errors even at age 4 or 5.

It is clear that morphological development does not follow a strictly U-shaped learning trajectory, which has often been discussed in relation to English past tense acquisition, in which it is proposed that children start out with rote-learned target-like forms (e.g. *ran, jumped*), then start to overgeneralize the most regular pattern(s) to irregular stems (e.g. **run(n)-ed*), and finally 'relearn' the irregular forms (see Marcus *et al.* 1992). Such a development predicts errors, especially in the second phase, and predicts that errors become rare when the child has learned the 'past tense rule'. Instead, errors are found to occur throughout development and may change in nature over time. And yet errors overall are rare – another counter to the U-shaped trajectory notion.

Similar data are obtained from Dutch, Danish and Hebrew, where plural marking depends on the phonology and syllable structure of the noun root as well as on gender, when the language has gender marking (see Ravid *et al.* 2008). Children seem to learn the likelihood with which a particular type of noun stem is associated with a particular affix, and their errors typically occur within the probability space of their target languages. In an analysis of child-directed speech (CDS) in these languages, Ravid and colleagues also found that the words and inflected forms addressed to children show the main tendencies of the target language in a clearer fashion than shown in a large corpus of Dutch adult-directed speech (ADS). The predictability of the plural allomorph from the end of the noun stem, the rhyme, was significantly better than in ADS. They conclude that CDS provides the child with a simplified, less complex version of the

grammatical system, since the regularities are overrepresented and the more exceptional patterns are absent or low frequency. For example, a larger proportion of the nouns in CDS end in schwa, and these nouns predictably take the *-n* plural. This group also includes diminutives ending in *-tje* which constitute 5 per cent of CDS, but only 1 per cent of ADS. In contrast, ADS has more nouns ending in full vowels. The nouns are typically of Latinate origin and are less predictable than the nouns ending in schwa (Ravid *et al.* 2008: 56–59).

In sum, the plural is a grammatical category that is conceptually relatively simple and acquired early, but may be structurally complex because it is encoded by several allomorphs that may be conditioned by including phonotactics and gender. Children segment the plural allomorphs readily, but take some time to associate the appropriate ending with the noun stem properties.

12.5.2 Tense–aspect marking

For languages with marking of tense and aspect, the encoding of these categories is more involved than mapping the morphological encoding of plural as children have to develop the concepts as well as the encoding for all categories relevant in their target language. Tense markers locate an event on the time axis, as prior to, simultaneous with or posterior to speech time (e.g. *he walk-ed, he walk-s, he will walk* in English, where some tense forms are formed by affixation, and others by auxiliaries in combination with the infinitive or past participle of the main verb). But it is also possible to relate events from a reference point that is not speech time (e.g. *when John was in Paris, he walked to the Eiffel Tower*) (see Comrie 1985). Aspect markers, in contrast, characterize the temporal contour of the event (Comrie 1976). The English progressive, for example, encodes events as ongoing and incomplete (imperfective), as in *While John was playing, his dad cooked dinner*. A typical aspectual distinction is the one between perfective and imperfective, as found in Slavic languages (Stoll 2005). The exact meaning of tense and aspect markers is further determined by the lexical properties of the verb phrase that encodes the event. For example, activity verbs including *play* or *run* encode a temporally unbounded event. Verb phrases like *run a mile* or *find the ball* encode the right temporal boundary or end of the event, which can be punctual, as in *find the ball*, or durative, as in *run a mile*. Vendler (1967) distinguishes four such types of lexical aspect or Aktionsarten: temporally unbounded states (*be, know*) and events (*run, play*) and temporally bounded events which are punctual achievements (*find*) or durative accomplishments (*drown*). Semantically, lexical aspect interacts with grammatical aspect and tense in subtle ways. In English, states and achievements do not combine with the progressive (**He was knowing, *He was finding*). Also, when punctual events are combined with a durative marker, this does not denote that the event took longer, but that it

happened several times. Thus, *He sneezed for two hours* does not mean that there was one very long sneeze, but that there was an extended period with a series of sneezes. Morphologically, languages differ as to which temporal and aspectual notions they encode by grammatical categories, and in their exact semantic effects and constraints (Comrie 1976, 1985).

For the child, the acquisition of tense and aspect poses a manifold learning problem: learning tense categories requires the ability to abstract from the here and now, to recall past events from memory and to anticipate the future. To distinguish aspect categories, it is necessary to be able to categorize temporal properties of an event, e.g. whether it is bounded or unbounded, punctual or durative. That is, children not only have to acquire the form and function of tense and aspect markers along with the perspective taken on the event that is encoded. In addition, in order to mark the categories, the appropriate morphological forms and contexts of use need to be determined. Tense and aspect distinctions do not always form a full paradigm and do not always combine with all verbs. Languages differ as to whether they mark only tense (Hebrew) or aspect (Mandarin Chinese) or both (English, Russian), and as to how the categories are encoded: by verb inflection, separate particles, or stem modification (see Wagner 2012 for a survey).

Despite the conceptual complexity and the crosslinguistic diversity in encoding tense and aspect, the findings reported in the child language literature are remarkably similar: a strong association between perfectivity and early past tense markers is found in early first language acquisition (Li & Shirai 2000). In English, for example, early past tense markers are used with resultative verbs, e.g. *broken*, whereas the progressive is predominantly used with activity verbs, e.g. *riding, jumping, kicking.* Similar associations are observed in other languages, across typological divides, e.g. Finno-Ugrian and Slavic languages (Stoll 2005, Weist, Wysocka & Lyytinen 1991), and typologically diverse South-East Asian languages (Shirai 2014). While some authors have proposed that children's early use of past markers is restricted because they are cognitively unable to represent past event and see utterances like *the vase is broken* as the description of the current resultant state (Antinucci & Miller 1976, Bronckart & Sinclair 1973), there is evidence that young children at the onset of morphological development can represent and encode reference to past events, and also understand the respective linguistic markers in comprehension tests (Wagner 2012).

A second type of explanation is frequency- and semantics-based: the associations between tense–aspect marking and lexical aspect found in learner languages follow the most frequent patterns in the input language. Children first focus on high frequency and prototypical options and differentiate their tense–aspect system only later (Shirai 2014, Shirai & Andersen 1995, Wagner 2012). In this view, children start off with a probabilistically determined subset of options. But what determines the prototypicality of

certain tense and aspect markers for certain types of events? Wagner (2009) carried out a comprehension experiment with 4-year-olds and adults: participants were asked to match pictures with completed or incomplete actions to the respective tense or aspect marker (e.g. *painting a flower / painted a flower*). Adults were at ceiling in all cases that represented a prototypical event, e.g. for the sentence *The woman was painting a flower*, a picture where the agent (a woman painter) was present, and had slightly lower scores with less complete depictions of an event, e.g. for the sentence *Somebody was painting a picture*, a picture with an unfinished flower but no painter. Children also scored well on the prototypical associations, but were less successful in the other condition. Wagner (2009) concluded that with aspect markers that express an ongoing activity, children expect an agent. In case of telic events like *build a house*, they expect the house to be completed, not that there is just the intention to complete a half-finished house. This is in line with information-theoretic accounts that propose that processing is easiest when the linguistic markers correspond with the event realization (Bohnemeyer & Swift 2004). The link between event structure, information structure and processing ease can explain why we find such strong linguistic associations between tense–aspect morphology and certain event types crosslinguistically.

In sum, in the area of tense–aspect marking the semantics can become cognitively and conceptually quite intricate. However, language itself plays a role in shaping these concepts: languages provide different morphological cues, and they also differ in how fine-grained their grammatical categories are when it comes to carving up the semantic space. What is clear from the acquisition literature is that children are influenced by the input language from the earliest stages to determine the language-specific distinctions and context of use.

12.5.3   Competing cues in form–function mapping

Studies on the acquisition of morphology have shown that children do not acquire all aspects of a grammatical category at once, although they start learning and generalizing very early (Chapter 8). Evidence for the gradual course of acquisition comes from distributional analyses where it is shown that children learn inflectional paradigms in a piecemeal manner (see the productivity criteria discussed in Section 12.4.2), as well as from experimental studies where children fail to apply a marker in a particular condition.

But what makes some linguistic structures harder to acquire than others? The course of acquisition is determined by a number of processing factors that promote or inhibit learning. Based on an analysis of crosslinguistic differences in language structures and the course of acquisition, Slobin (1973, 1985a) formulated forty so-called operating principles that children were assumed to make use of when processing their target

language (see Slobin 1985a). These operating principles were assumed to guide children in their initial segmentation of speech (Peters 1985), and provide strategies and procedures for attending to and storing the input they receive, and for deriving linguistic patterns from it by organizing segmented elements into units and mapping them to their function (Slobin 1985a). For example, inflectional morphology often surfaces as suffixes, and thus children would pay attention to the end of a unit and store it separately from the unit (Slobin 1985a). Operating principles need to be flexible enough to allow the acquisition of typologically different languages and thus account for the variable mappings between different levels of representation attested in different languages. For example, children may apply a particular solution that has proved successful in similar situations, and this repetition leads to entrenchment. Strengthening can explain why gender is acquired early in Hebrew, where it is a pervasive category that is relevant for a wide range of agreement phenomena and why plural formation is acquired late in Romance languages, where it only plays a marginal role (Slobin 1985a).

But the concept of operating principles, although psychologically plausible, was criticized at the time. This was because the operating principles were designed to deal with the divergent processing problems different languages pose, and offered a large repertoire of strategies and learning procedures. There were a number of problems raised. First, the underlying assumption was that children's language development would follow the same cognitive development irrespective of the language being acquired. Subsequent typological research led to a rethinking of the relationship between thought and language in theories of linguistic relativity (Bowerman 1985, 1994, Bowerman & Choi 2003, Slobin 1997b, d), and in the acknowledgement that much of children's category formation is language specific.

Second, the operating principles were criticized as they lacked predictive power; there was no theory of which operating principle to apply as they are not mutually exclusive (Bowerman 1985, Hakuta 1988). This problem is addressed in recent frameworks that see acquisition as a process of integrating probabilistic cues. While individual grammatical categories within and across languages pose their own acquisition problems in terms of the exact form–function mappings they represent, they also share learnability issues that relate to general theories of information processing. From a processing perspective, each language and dialect has its own affordances regarding cues it offers for learning its structure. Bates and MacWhinney (1987) conceptualized these cues as being in competition with one another such that the ease or difficulty of acquisition can be predicted by the validity of a cue (a function of its availability and reliability) in relation to the cost of processing it. For example, the English plural suffix *-s* is highly available because it serves as the plural marker on

almost all nouns, but it is not very reliable because -s affixes also encode third-person singular agreement on verbs, and the genitive on nouns. In contrast, the progressive marker *-ing* is available and reliable, as it does not serve another function. Morphological markers that show a one-to-one mapping of form and function thus have the highest cue validity (e.g. the very regular agglutinative affixes found in Turkish); cues that have several functions have lower validity. Cue cost computes the difficulty of processing: affixes, for example, are easier to process than single-segment changes of the stem because extra syllables are easier to perceive than segment changes. Also, frequent elements are easier to notice and to memorize than infrequent ones.

The Competition Model allows prediction of the order of acquisition in different languages, and it is well suited for studying phenomena where certain functions are redundantly marked by grammatical markers and syntactic operations. Regarding syntactic transitivity, for example, word order is a very reliable cue in English, but less reliable in German, where subjects and direct objects can be reordered because their grammatical status is encoded by case markers (*Die Kuh sieht der Mann* 'The cow-ACC sees the man-NOM'). In German, case marking on the definite article is the most reliable cue, but not always available: only the masculine nouns have a unique and distinct morpheme for the dative definite article, and not all nouns are masculines, and not all masculines are accompanied by the definite article. A study on German revealed that young children rely on word order, and only older children are able to process case cues (Lindner 2003; see also Chan, Lieven & Tomasello 2009, Dittmar, Abbot-Smith, Lieven & Tomasello 2008a, Kempe & MacWhinney 1998).

In the Competition Model the effect of the 'weight' of different cues on acquisition is computed. Children rely on different cues in different languages depending on the coalition of cues available: e.g. in the interpretation of transitive sentences that encode 'who is doing what to whom' (e.g. *The dogs the eraser push*), English children rely heavily on word order and make little use of cues like animacy or agreement. In Italian, however, children initially rely on animacy, before they learn to attend to agreement. Word order is not a cue they use much because word order is more variable in Italian (MacWhinney, Bates & Kliegl 1984). These early cross-linguistic studies show that children start out by paying attention to strong and reliable cues, and that processing may change as other aspects of the system are acquired. Currently, these ideas are being developed into more general processing theories that account for first and later second language learning (MacWhinney 2012). Research on first language acquisition has shown how children tune in to the grammatical properties of their language. This 'learned attention' (Ellis 2006a,b) in turn influences how they initially process additional languages. Exposure to an analysis of the input language leads to perceptual fine-tuning that in turn affects what can be processed by the learner. Some cues may initially escape perception

and processing because they are low in perceptual salience and/or reliability (see Ellis 2006a, b, MacWhinney 2012).

12.6   Summary and outlook

This chapter has focused on generalization processes as well as theories and methods to account for the developmental processes observed in the acquisition of grammatical categories. Analysis of children's spontaneous speech or systematically elicited data has informed us about the time course and nature of generalization, as well as about individual differences in children's learning trajectories. Crosslinguistic comparisons have led to deeper insights into the factors that influence acquisition, for example, the five volumes edited by Slobin (1985–97), which present detailed descriptions of the acquisition of morphosyntax in a wide range of typologically different languages; the chapters in Berman and Slobin (1994), which systematically compare data from narratives, as do chapters in Strömqvist and Verhoeven (2004); and the chapters in Bavin and Stoll (2013), which discuss the acquisition of ergative morphology in a variety of languages.

Experimental work has also provided evidence of children's generalization processes and computational modelling has tested their empirical likelihood. However, while these models are very successful in dealing with the form side of language, they are currently less well suited for integrating semantic and pragmatic factors, simply because these are harder to operationalize in experimental designs, because of their context-dependency.

Morphological categories pose a language-specific acquisition problem that involves the integration of multiple cues and domains (phonology, prosody, semantics, pragmatics, syntax). For example, the choice of plural allomorphs depends on phonetic and prosodic factors, as well as morphological categories such as gender. While there is a substantial body of research showing how children build the paradigm of a particular inflectional category and what type of errors they make, less is known about individual differences. Most production data still come from rather small datasets involving very few children, where developmental processes are studied in large time-frames with relatively few data points. In order to unravel the fine-grained interaction that contributes to the form–function correspondence in grammatical categories especially for lower-frequency phenomena, the richness of the existing datasets needs to be exploited.

This is particularly relevant when looking at older children. Young children seem to acquire the basic grammatical properties of their target language with relative consistency, and a large part of the acquisition literature focuses on categories attested in many languages (like number, tense, aspect, case). However, languages often also have smaller 'niches' of

inflectional paradigms that do not extend to the whole category. For example, in German noun morphology distinguishes four cases, but they are morphologically distinct only in the masculine singular (see above), and even this distinction is hard to perceive in spoken language, and absent in some varieties where phonologically reduced forms of the determiner are used. Hence, dative marking is difficult for German children (Behrens 2011, Szagun 2004) and even 7- or 8-year-old children may have difficulties in processing it, in particular if they have little exposure to the standard variety. However, command of case is relevant for academic language at school when children are required to show full command of case marking in writing (see Chapter 21). That is, individual differences regarding the general degree of linguistic productivity and the command of fine-grained grammatical detail may become more pronounced as children grow older (see Hart & Risley 1995 for a longitudinal study on the effect of the quality and quantity of the input on lexical and grammatical development and later school success).

Suggestions for further reading

Berman, R. A., & Slobin, D. I. (eds.) (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, NJ: Erlbaum.

Bowerman, M., Brown, P., Eisenbeiss, S., Narasimhan, B., & Slobin, D. I. (2011). Putting things in places: Developmental consequences of linguistic typology. In J. Bohnemeyer & E. Pederson (eds.), *Event Representation in Language and Cognition* (pp. 134–65). Cambridge University Press.

Slobin, D. I. (ed.) (1985–97). *The Crosslinguistic Study of Language Acquisition*, vols. 1–5. Mahwah, NJ: Erlbaum.

Wagner, L. (2012). First language acquisition. In R. Binnick (ed.), *The Oxford Handbook of Tense and Aspect* (pp. 458–80). Oxford University Press.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A.,. . . & Dressler, W. U. (2011). On the role of morphological richness in the early development of noun and verb inflection. *First Language*, 31(4), 461–79.