REFERENCES

Bauer, R. S. 2006. The stratificaton of English loanwords in Cantonese. *Journal of Chinese Linguistics*. 34 (2): 172–191.

Bauer, R. S., and P. K. Benedict. 1997. *Modern Cantonese Phonology*（摩登廣州話語音學）Berlin: Mouton de Gruyter.

Bauer, R. S.（包睿舜）, and K. H. Cheung（張群顯）. 2005. *Database of Cantonese Words Derived Through Diminutive Bianyin*（粵語小稱變音詞語派生的詞彙）, 載鄧景濱主編 [Jingbin, Deng. ed.],《第九屆國際粵方言研討會論文集》*Proceedings from the 9th International Yue Dialects Conference* (Macau, 2003). 澳門：澳門中國語文學會 [Macau: Macau Chinese Language Association]. 174–179.

Bauer, R. S.（包睿舜）, and Cathy S. P. Wong（黃倩萍）. 2010. New loanword rimes and syllables in Hong Kong Cantonese. In *The Joy of Research II, A Festschrift in Honor of Professor William S-Y. Wang on His Seventy-fifth Birthday*《研究之樂，慶祝王士元先生75壽辰學術論文集》, eds. PAN Wuyun 潘悟云 and SHEN Zhongwei 沈鐘偉, 1–24. Shanghai: Shanghai Education Publishing（上海：上海教育出版社）.

McCoy, J. 1980. The reconstruction of upper register nasals and laterals in proto-Cantonese. In *Contributions to Historical Linguistics, Issues and Materials,* eds. Frans van Coetsem and Linda R. Waugh, 200–213. Leiden: E. J. Brill.

Xianggang Yuanxue Xuehui 香港語言學會 (Linguistic Society of Hong Kong). 2002. *Guide to LSHK Cantonese Romanization of Chinese Characters*《粵語拼音自表》. 香港：香港語言學學會.

Zhan Bohui（詹伯慧）, and Risheng Zhang（張日昇）. 1987. *A Survey of Dialects in the Pearl River Delta, Volume 1, Comparative Morpheme-Syllabary*《珠江三角洲方言字音對照》. Hong Kong: New Century Publishing House（香港：新世紀出版社）.

4

Linguistic Adaptation: The Trade-Off between Case Marking and Fixed Word Orders in Germanic and Romance Languages

Christian BENTZ
Department of Theoretical and Applied Linguistics, University of Cambridge

Morten H. CHRISTIANSEN
Department of Psychology, Cornell University

Foreword

Understanding language evolution in terms of cultural transmission across generations of language users raises the possibility that some of the processes that have shaped language evolution can also be observed in historical language change. In this paper, we explore how constraints on production may affect the cultural evolution of language by analyzing the emergence of the Germanic and Romance languages from their proto-languages. Specifically, we focus on the change from flexible but OV (Object-Verb) dominant word orders with complex case marking to fixed SVO (Subject-Verb-Object) word order with little or no noun inflection in modern Romance and Germanic languages. We suggest that constraints on second language learners' ability to produce sentences may help explain this trade-off between nominal case marking and fixed word orders. Furthermore we assume that this scenario might not only hold for the Germanic and Romance branches of the Indo-European languages but could perhaps be extended to a variety of language families and areas. We conclude that historical data on linguistic change can provide a useful source of information relevant to investigating the cognitive constraints that affect the cultural evolution of language.

1.   Introduction

If language has evolved primarily through cultural transmission (e.g., Christiansen and Chater 2008), then language evolution and language change may not be clearly distinct in a theoretical sense (Christiansen, in press). Rather, it may be expected that the processes proposed to underlie patterns of historical language change (e.g., grammaticalization) also have been at play across the longer timescale of language evolution (Heine and Kuteva 2009). Thus, diachronic change may be construed as a microcosm of language evolution and potentially provide a rich source of data to illuminate potential constraints on linguistic adaptation.

In this chapter we want to address the following question: Are there diachronic data on language change that indicate that constraints on human cognition have shaped language on a historical time scale? To answer this question, we consider as a case study the pathways from a reconstructed Indo-European proto-language to the Germanic and Romance languages of today, focusing on how limitations on production may affect linguistic adaptation. We sketch an account that offers production as one of the multiple cognitive constraints explaining historical language change. This account highlights the sequencing problems that a second language (L2) learner faces when producing a sentence. Together, these observations corroborate our suggestion that historical language change may be construed as *linguistic adaptation*, i.e., language structures adapt to domain-general constraints deriving from the human brain, rather than vice versa.

2.   The Diachronic Pathways

2.1   From Proto-Indo-European to Proto-Germanic and Classical Latin

There are mainly two tendencies prominent in diachronic accounts of the change from the reconstructed Indo-European proto-language to Modern Romance and Germanic languages. First, the canonical word order of transitive sentences seems to have changed from free word order (perhaps with slight OV dominance) to fixed SVO patterns. Second, while Proto-Indo-European used a complex set of nominal case marking paradigms to indicate who-did-what-to-whom, most of its modern linguistic descendants no longer adopt this strategy. We can follow this pathway of case marking decline in Romance and Germanic languages as far back as to the period of roughly 2000 BC. By comparing case marking systems in attested languages such as Classical Latin, Old Greek, Hittite and Sanskrit researchers have come to the conclusion that Proto-Indo-European had a fully-fledged system of nominal case affixation with eight (Behaghel 1923; Meier-Brügger and Krahe 2002; Hutterer 2002) or even nine paradigms (Speyer 2007): Nominative, Genitive, Dative, Accusative, Vocative, Instrumental, Locative, Ablative and Allative. The issue of a canonical versus flexible word order, however, is a rather controversial one. Ebert (1978) and Speyer (2007) argue that the main clause in Proto-Indo-European must have been dominated by the SOV type, whereas Admoni (1990) highlights the general flexibility

of the word order, which goes hand in hand with the nominal marking strategy. As we will see below, the issue of a prevailing word order is hard to settle even when corpora are available. Therefore we do not dwell on this issue here but adopt the compromise assumption that the Indo-European proto-language most likely displayed flexible word order patterns, perhaps with the tendency towards prevailing SOV patterns. Henceforth, we categorize word orders by using a notation similar to the one introduced by Van Everbroeck (2003). According to this notation the Proto-Indo-European type would be XXX in reference to Admoni (1990) or SOV in reference to Speyer (2007) and Ebert (1978). The three letters indicate the assumed word order (SOV, SVO, etc., or XXX for flexible word order). Using these definitions, we categorize the Proto-Indo-European language as XXX (SOV). The SOV in parentheses indicates that there might have been a distributional tendency towards using SOV structures, albeit in a generally very flexible system.

The change from Proto-Indo-European towards Germanic, and Westgermanic in particular (spoken in northern Germany and Scandinavia), on the one hand, and from Proto-Indo-European to Classical Latin, on the other hand, took about 2000 years from ~3000BC to ~1000BC/500BC (Speyer 2007, 13). In this time span a simplification of the nominal marking system has already taken place in both language families. For Germanic, Behagel (1923, 477) argues that only the ablative case was lost in Proto-Germanic and replaced by dative markers. However, it seems likely that the locative functions also started to get covered by the dative paradigm (Hutterer 2002, 54). In this context, Speyer (2007, 72) concludes that the reduction from eight or nine cases in Indo-European to six in Proto-Germanic is an essential part of the so-called 'simplification of the Proto-Germanic nominal system'. For Latin it is clear that the Allative and Instrumental cases were lost, which leaves at least five and at most seven paradigms: Nominative, Accusative, Dative, Genitive, Ablative as well as Vocative and Locative (which were also already in decline).

What about the canonical word order of these periods? With reference to Gothic rune inscriptions and the Old English text *Beowulf*, Ebert (1978, 35) assumes that verb final structures must have gained dominance already in Germanic times. Again, because of the lack of direct corpora this is a controversial issue. The use of six distinct case marking paradigms in Germanic would have allowed speakers to change word orders without risking the loss of relevant information (Hutterer 2002, 64). All in all, the available evidence suggests that the trend towards a SOV dominated system was realized more strongly in Proto-Germanic than in Proto-Indo-European, although using flexible word order patterns was still possible because of the six productive case paradigms. Therefore, we would categorize Germanic as SOV (XXX) language.

For Classical Latin, there is a fair amount of corpora available. As will be shown in the next section in more detail, the Classical Latin system can also be categorized as SOV (XXX), since the SOV pattern was used on a regular basis, although the case system was still productive enough to allow flexible use of word order patterns. Thus, the changes in the branches from Proto-Indo-European to Proto-Germanic as well as to Classical Latin seem to follow similar pathways. In the roughly two thousand years between these stages both the Germanic and the Latin language seemingly strengthened SOV as canonical word order and both lost two or three productive case-paradigms (Allative, Locative and Ablative in the Germanic case; Allative and Instrumental in the Latin case). But from this point (~500BC) onwards, Germanic and Latin, as well as the languages splitting off from these, started to take different paths in terms of their

strategies for encoding thematic information. In the following, we propose that one key factor in determining the paths toward further case loss and canonical word order was the strength of influence of production constraints from L2 speakers. Therefore, we are concentrating on the developments from Classical Latin to the Romance languages and from Germanic to Modern High German and Modern English as examples of *linguistic adaptation*: here, the shaping of grammatical structures by the cognitive constraints of the speaker population.

2.2 Latin to the Romance languages

Taking the development of Latin towards modern Romance languages as an example of linguistic adaptation, we concentrate on simple transitive sentences because they can be considered the neutral prototype of other more complicated constructions (Slobin and Bever 1982). There are two interesting changes to this sentence type occurring in the time span between Latin (~500 BC – AD 500) and recent Romance languages:

(i)   While Latin had a five to seven case system (i.e., Nominative, Accusative, Dative, Genitive, Ablative, (Locative), (Vocative)) all subsequent Romance languages use fewer cases.

(ii)   The word order in simple transitive sentences has changed from OV (foremost realized in SOV and OSV) to SVO.

Consider for example the following aphorism by Vergil:

(1)   *Fata*   *viam*   *invenient*
      fate-NOM-PL   way-ACC-SG   find-3P-PL-PRE-ACT

(2)   *I fati trovano una via* (direct Italian translation)

(3)   *The fates find their way* (direct English translation)

Latin makes use of the accusative marker to indicate who finds whom: fata via-m, but in Italian (as in English) the marker has vanished and the problem of assigning thematic roles is solved by using a strict SVO word order. The nature of the change in word order has been the subject of some debate among specialists of Romance languages (Pinkster 1991; Lee 2002; Salvi 2004). We therefore tabulated the number of sentences with different {S,O,V} ordering in simple declarative sentences. Using the two complete sets of counts from the classical period (Caesar and Petronius) and the later *Peregrinatio* (AD 400) from Pinkster (1991), we obtained the distribution shown in Figure 4.1. As Pinkster notes, S preferably takes initial position and O precedes V more often than the other way around. This displays the OV pattern as predominant, albeit in a flexible system.

In contrast, Modern Romance languages are widely assumed to have a clear predominance of SVO word order (Harris 1988; Lee 2000; Salvi 2004). For example, Slobin and Bever (1982) report word order frequency data for Italian indicating a clear predominance of SVO sentences

Figure 4.1   The frequencies of different word orders in Latin (based on Pinkster 1991, 72)
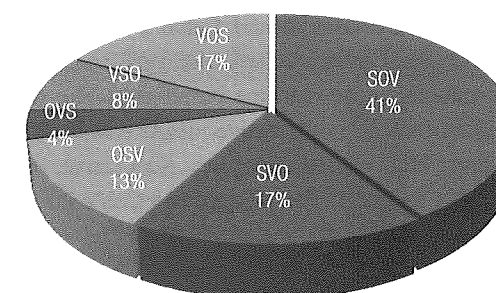
(adults: 82% SVO, 2% SOV, 0 % OSV; children: 72% SVO, 1% SOV, 1 % OSV). Thus, usage of the OV patterns has declined to a minimum in Romance languages, such as Italian.

To summarize, while Classical Latin is an example of a SOV (XXX) system, displaying a strong tendency to use SOV patterns and additional nominal affixes, the modern Romance languages can be categorized as SVO languages, i.e., fixed SVO word order systems without nominal case marking (except Romanian, which preserved nominal markers for the opposition Nominative/Accusative and Genitive/Dative). As outlined by Herman and Wright (2000), the Vulgar Latin period (~100BC–~500AD) can be seen as a gradual transition from Classical Latin to the modern Romance type. Vulgar Latin can be defined as SOV (SVO), because SOV was still predominant although SVO patterns already started to become more frequent while nominal case marking started to decline. The nominative, accusative and ablative merged into one category because of semantic confusions and phonological leveling. Likewise, the genitive and dative merged to become an oblique case with a separate marker in contrast to the nominative (Herman and Wright 2000, 57). Depending on regional differences, this two-way distinction was either further weakened, as in Italian, Spanish and French, or preserved as in the Romanian case (Herman and Wright 2000, 58). However, the overall tendency from Classical Latin towards the Romance languages is relatively clear: While case declensions started to become eroded, the SOV dominant pattern changed into a fixed SVO order. As will be shown in the next section, these tendencies are also reflected to different degrees in the development from Proto-Germanic to Modern English and Modern High German.

## 2.3 Germanic and the contrast between Modern English and Modern High German

The changes in the Germanic branch are quite similar to the ones observed in the Romance example. Deriving from Proto-Germanic, the split-off point between the German and English branches directly precedes the Old English (~450–1150 AD) and Old High German (~750–1050 AD) stages (Baugh and Cable 2006, 52). Thus, it is not surprising that OE and OHG were still quite similar in terms of word order and number of case marking paradigms. Analyzing the

available corpora we find the early tendency to use the so-called *Verbzweitstellung*, i.e., the verb in second position in declarative main clauses in both OHG (Admoni 1990; Ebert 1978; Speyer 2007; Braune et al. 2004) and in OE (Baugh and Cable 2006; Speyer 2007; Hogg and Blake 1992). This gradual change of word order towards the *Verbzweitstellung* seems to go hand in hand with further loss of productive case marking paradigms. The replacement of the vocative by the nominative and the instrumental by the dative means that both OE and OHG are left with the same four case paradigms: Nominative, Genitive, Accusative, and Dative (Admoni 1990; Behaghel 1923; Hutterer 2002; Campbell 1959; Baugh and Cable 2006).

However, from this point onwards the changes in the English branch are very different from the ones in the German line. While the Middle English period (~1150–1500AD) is known as the period of 'leveled inflections' (Baugh and Cable 2006, 52), the German nominal system of affixes has been more strongly preserved until today (Dal 1962, 4). A closer look at the changes in the case marking paradigms reveals that the pathway from Old High German to Middle High German and to Modern High German is as well characterized by analogical leveling of formerly differentiated case declensions, which is, however, not nearly as radical as in the Middle English case. Except for some personal pronouns like *him* or *whom* in Modern English there is no trace left of the Old English four case paradigms, which are still reflected in German nominal affixes, pronouns and articles. Although Allen (1997, 2005) argues that if different writing habits of OE and ME scribes are taken into account, the transition from OE to ME seems less abrupt, this does not change the big picture that loss of case inflections was noticeably intensified in the few hundred years between ~1150 and ~1500. In Section 4, we seek to explain this phenomenon with reference to L2 production constraints. But for now it is important to note that the Middle English period might be categorized as SVO (SOV) type, and the Modern English period as fixed SVO. In Modern High German the second position of the verb is dominant in simple transitive main clauses, whereas subordinate clauses still regularly display the 'old' verb final (VF) pattern, which was still grammatical in main clauses of Middle High German. Therefore the German categories are SVO (SOV) for both Middle High German and Modern High German.

In summary, we have briefly outlined the emergence of the Romance languages as well as Modern English and Modern High German from their Classical Latin and Proto-Germanic predecessors. Figure 4.2 provides an overview of all the changes in the different branches discussed here. Note that all the categories reflecting number of nominal case marking paradigms, word orders and dependent marking are still crude. But we submit that the general trends from Proto-Indo-European to Germanic and Latin as well as their modern descendants are captured in this outline. If a free word order (XXX) with SOV tendencies is assumed for the putative Indo-European proto-language, then it is remarkable that the Germanic and Romance languages all tend to shift this word order towards a fixed SVO pattern. A second trend that can be deduced from this illustration is the loss of nominal case marking paradigms throughout the course of historical language change. More precisely, it is interesting that the rate of case loss seems to differ quite substantially between different periods and languages. While only 2–3 paradigms were lost during the about 2000 years between Proto-Indo-European and Proto-Germanic as well as Classical Latin, some 4–5 case paradigms almost completely vanished within just a few hundred years during the periods of Vulgar Latin and Middle English. This suggests that either the crucial factor governing case loss was strongest in these periods, or that

Figure 4.2 The diachronic pathways of nominal case marking paradigms and word order in German, English and the Romance languages
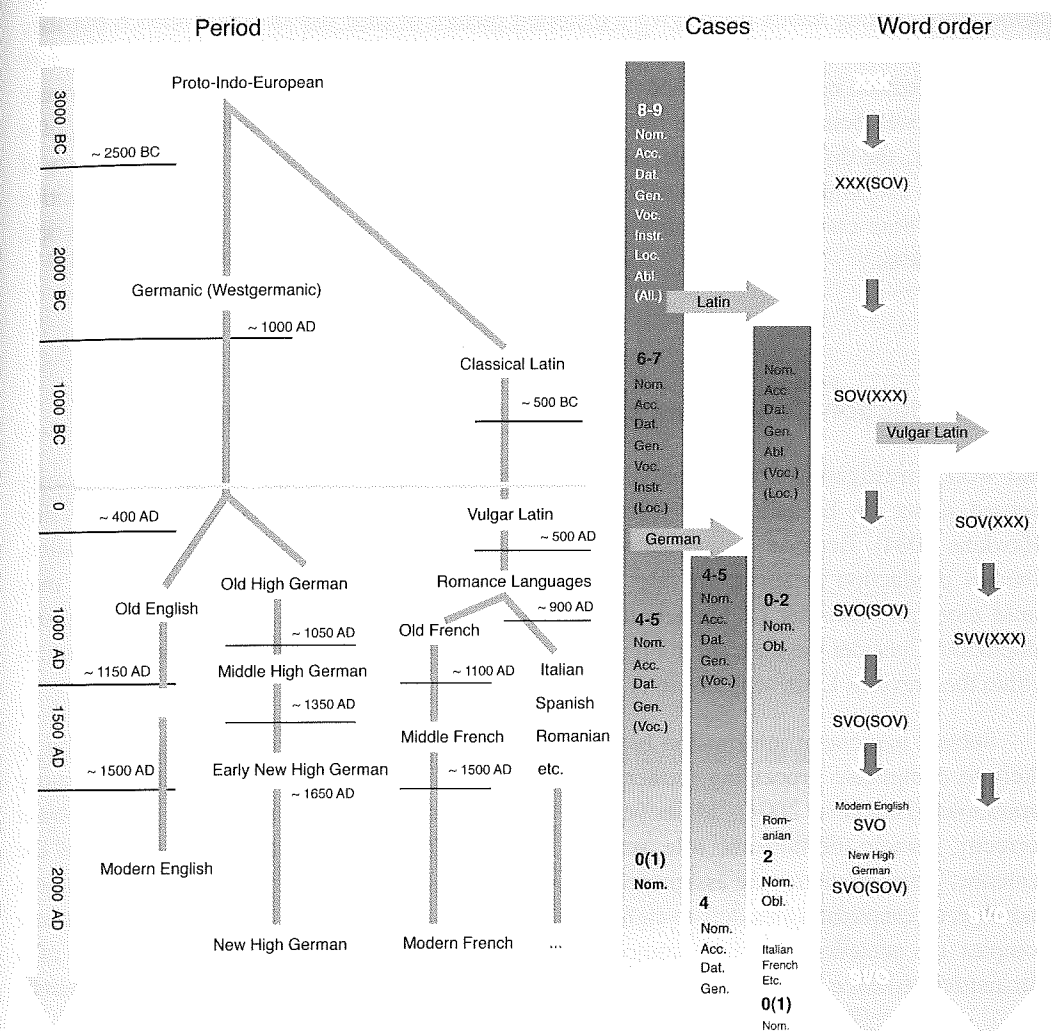
there are various factors accounting for a 'baseline' of case erosion, which were reinforced by an additional factor in Middle English and Vulgar Latin. We propose that the impact of L2 learners on nominal paradigms is one of these relevant driving forces. In Section 4, we provide more detailed information about the social backgrounds of Vulgar Latin and Middle English to corroborate this view.

However, we are not suggesting that L2 learning constraints are the only factor relevant for the loss of case paradigms. There are likely to be a multitude of intertwined social, cognitive and linguistic factors accounting for these tendencies in general. For both the Romance and Germanic branches, it has been argued that different writing styles of scribes in earlier times

complicate the picture of the correct usage of nominal affixes and canonical word orders, let alone the differences between written and spoken language in the same period (Pinkster 1991; Allen 1997). Still, we contend that the overall tendencies must be connected to an additional factor, since writing habits might give a distorted picture of case use in actual speech but they do not account for the functional preservation or loss of case categories per se (for a more detailed discussion of 'linguistic' and 'extra-linguistic' factors, see Bentz and Winter, in press). The factors we focus on here are the production constraints facing non-native speakers of a language, i.e., the difficulties in determining the appropriate nominal affixes that late learners of Vulgar Latin and Middle English would have experienced when producing even simple transitive sentences. These constraints are discussed in the next section using examples from Classical Latin and Italian.

3. Production Constraints as a Source of Language Change

Past work investigating how cognitive constraints may shape language change has primarily focused either on limitations on learning (Polinksy and Van Everbroeck 2003) or parsing (Hawkins 2004). Because comprehension can be managed by integrating partial information, whereas production means specifying the complete sequence, we suggest that the latter may cause more problems for L2 learners and therefore provide a strong pressure toward linguistic adaptation. Consider the diagram in Figure 4.3, illustrating the complex dependency relationships within the previous Latin sentence in (1).

Subject agreement information has to 'bypass' the direct object to get to the verb. This is likely to complicate processing further in sentences with embedded structures due to memory limitations (Hawkins, 2004). Moreover, the information required to inflect the direct object correctly, namely the thematic role assigned by the verb, is not provided until the end of the sentence. Thus, thematic role assignment has to be 'back-projected' from the verb to the subject and object, complicating the left-to-right sequencing of words in language production. The more complex the sentence, the more complex the role assignment becomes. In the example of a ditransitive sentence in (4), the speaker has to assign three roles and therefore inflect two nouns:

(4) *Magister*   *puell-ae*   *libr-um*   *dat*
   teacher-NOM-SG   girl-DAT-SG   book-ACC-SG   give-3P-PRE

This complexity contrasts with the much simpler set of dependency relationships shown in Figure 4.4 for the Italian transitive sentence in (2) (which would also hold for the English sentence in (3)). Crucially, all arrows proceed from left to right, except the one assigning the thematic role of agent to the subject (mapped onto the voice character of the verb). But as the subject does not inflect according to the thematic role in Italian (at least for proper nouns) this is not a problem.

Thus, Italian (and English) SVO word orders fit well with a simple left-to-right sequence production mechanism. Obviously there is a trade-off between two constraints within such

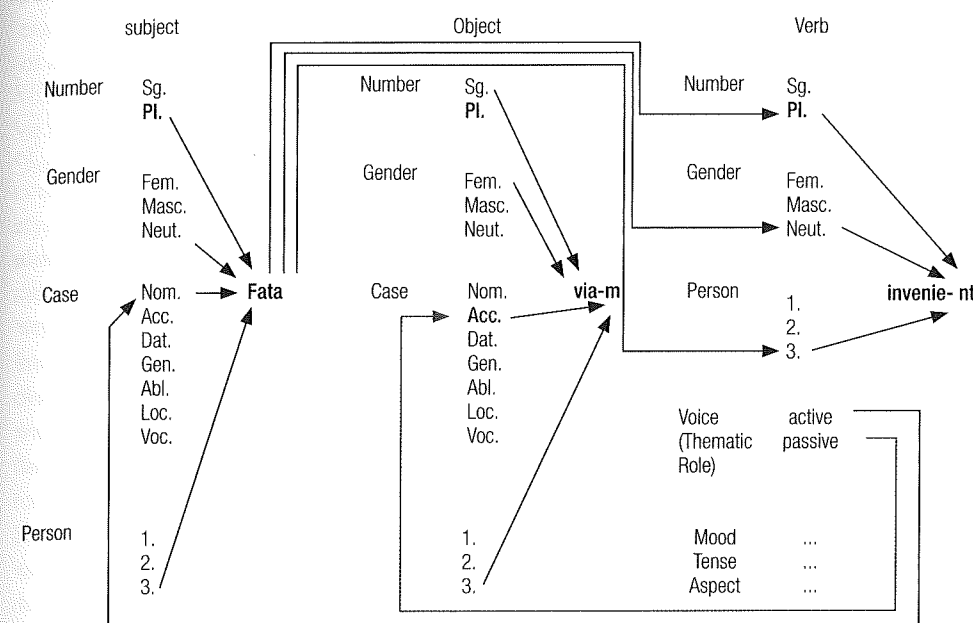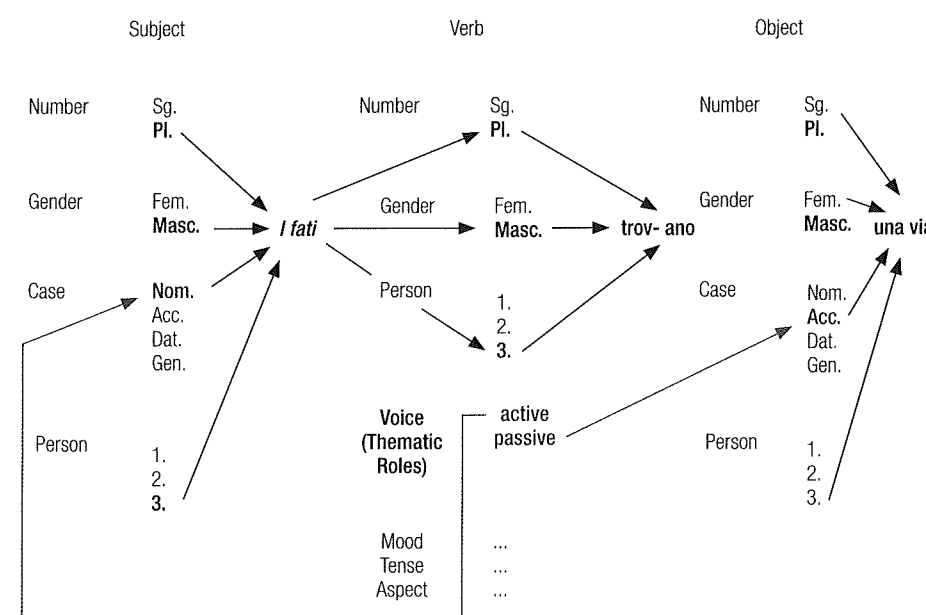Figure 4.3   SOV with case marking in the Latin sentence Fata viam invenient

Figure 4.4   SVO without case marking in the Italian sentence *I fati trovano una via*

simple transitive sentences. On the one hand, the verb should follow the subject because then the information regarding agreement of person and number is available when the verb has to be inflected. On the other hand, the verb should precede both subject and object to facilitate case marking. Given these constraints, the change from Latin OV and case marking of proper nouns to Italian SVO and no case marking makes sense given a production system that generates left-to-right sequences. Note, however, that based on the historical texts of Vulgar Latin it is difficult to determine whether the change of word order preceded or succeeded the loss of case paradigms, since both phenomena are intertwined and gradual in nature.

The production constraints outlined here are, in principle, the same for Middle English and Middle High German. There seems to be a trade-off between systems in which thematic information about the participants of a scene is reflected by nominal affixes—allowing a free word order—and systems in which this information is encoded using a canonical fixed word order as in the Italian example in Figure 4.4. The templates in Figure 4.3 and 4.4 illustrate how a fixed SVO word order without nominal case affixes is easier to handle from the perspective of left-to-right production because a) the subject precedes the verb which then can be inflected according to person and number, and b) there is no relevant information that has to be 'back-projected' by way of a nominal marker preceding the verb. This is in line with the assumption that word order and case marking are not independent structural features of languages but are rather intertwined from a typological perspective (Greenberg 1966, Universal 41). We will get back to this point in section 5. In the next section we first want to consider the social setting in which production constraints on case marking may come into play.

4. Meeting the Needs of Adult L2 Learners

Native speakers of Latin and Germanic would, of course, have been able to learn, process, and produce constructions such as (1) and (4) despite their complex dependency relationships, just as children are able to understand *who-did-what-to-whom* in Turkish, another heavily case-marked language with flexible but OV-biased word order (Slobin and Bever 1982). Therefore we suggest that an important pressure toward the simpler dependency relationships found in the Romance and Germanic languages came primarily from adult L2 learners, and only to a smaller extent from L1 acquisition.

As the Roman Empire grew, (Vulgar) Latin became its *lingua franca* and thus 'recruited' large numbers of non-native speakers. In a similar fashion, the English branch of the Germanic family was influenced by the presence of Scandinavian and Norman conquerors during several centuries of the Late Old English and Middle English period. These situations of intergroup contact may be seen as a large-scale historical parallel to the change from *esoteric* to *exoteric* communication, described by Wray and Grace (2007): Whereas the former is shaped by children's learning abilities, allowing the existence of idiosyncratic regularities that are hard for adult learners to master, the latter is tailored to the need for cross-group interactions, oftentimes by adult L2 learners. Thus, having to produce sentences with SOV word order and complex case markings as sequential *output* would have created considerable difficulties for adult L2 learners

of Latin and Old English. These difficulties provided an important pressure towards the SVO without case marking system in Modern Romance languages and Modern English.

But is there historical evidence that the 'recruitment' of non-native speakers might have impacted the structure of these languages? For the Vulgar Latin case, Herman and Wright (2000) describe the speech community between 100 BC and 500 AD, suggesting that speakers of other languages (e.g., slaves, merchants, inhabitants of the Romanized provinces) were continuously integrated into the wider Latin speech community on a large scale. This led to the atypical situation in which non-native L2 learners in many geographical areas outnumbered native speakers of Latin. Based on a detailed analysis of changes to Latin's formerly rich case system, Herman and Wright argue that the large amount of L2 speakers is likely to have shaped Vulgar Latin both in terms of morphology and syntax. The overall result would have been an increasing number of semantic and functional confusions between cases that previously had been distinctive: Ablative constructions were replaced by nouns with accusative markers and dative was used with prepositions to indicate possession instead of the classical genitive. Importantly, for our purposes, Herman and Wright note that (2000, 54), 'The accusative was originally used for the direct object of a transitive verb, and transitivity itself increased. Many verbs in Classical Latin were followed by a noun in the genitive, dative, or ablative case, but in Vulgar texts these verbs tend to take an accusative.' Because the word order in the period of Vulgar Latin still displayed mainly OV patterns, the tendency to over-generalize accusative case may be seen as a consequence of the difficulty of 'back projecting' thematic roles outlined in Figure 4.2. As a consequence of this ambiguous use of the case markers, the full system could no longer be maintained, and it shrank to a minimum. Therefore another strategy for solving who-did-what-to-whom dependencies was needed and emerged in the form of a fixed SVO word order.

In the Middle English period we find similar evidence for 'recruitment' of large amounts of non-native speakers. The first event of relevance in this context is the ongoing spreading of Danish and Scandinavian tribes to the English mainland between ~800 and ~1100 AD, i.e., in the Late Old English and Early Middle English period. After invasions and plundering of smaller villages, these foreign groups started to settle down extensively when Danish and Norwegian troops formed an alliance and gained the upper hand on the island. From 1014 to ~1029 England was reigned by kings of Danish descent. Even today the large-scale settlement of Scandinavian tribes is reflected in more than 1,400 Scandinavian place names in the United Kingdom (Baugh and Cable 2000, 95). A second and even more important phase of language contact started with the Norman Conquest in 1066. Baugh and Cable (2006, 108) consider this invasion as the immediate starting point of the inflectional erosion in Middle English. Although there are no concrete numbers of French people learning English as their L2 given for this period, we can get an impression of the extent of foreign speakers by the following quote: 'Among those of lower rank, whose position brought them into contact with both the upper and the lower class—stewards and bailiffs, for example—[…], the ability to speak English as well as French must have been quite general. […] The conclusion that seems to be justified by the somewhat scanty facts we have to go on in this period is that a knowledge of English was not uncommon at the end of the twelfth century among those who habitually used French; […]' (Baugh and Cable 2006, 123).

In short, both Vulgar Latin and Middle English seem to have recruited large numbers of L2 speakers by means of geographical spreading in the former case and invasion to the mainland in the latter. We suggest that these social factors led to substantial confusion between formerly distinctive case declensions and consequently to a stronger trade-off between SOV (XXX) and nominal case marking systems and the fixed SVO type without marking, when compared to other languages in the same branches such as German.

5. Possible Effects of L2 Acquisition beyond the Romance Languages

The claim that fixed SVO word order without case marking should be easier to use by L2 learners than flexible OV word order with case marking may appear problematic when compared to the typological frequencies of the world's languages. Standard typological analyses in terms of number of languages associated with certain word orders indicate that SOV word order is predominant: SOV 497; SVO 435; VSO 85; VOS 26; OVS 9; OSV 4 (Haspelmath et al. 2005, 330). However, if we look at the number of speakers that each language has, then a different picture emerges. Figure 4.5 shows the number of speakers for the twenty most frequently spoken languages in the world (*SIL Ethnologue online version*) and their respective word order according to the online version of WALS (World Atlas of Language Structures, Haspelmath et al. 2005). Adding up the numbers of speakers of these languages, a different pattern emerges: roughly 2,390 million speakers of SVO languages against 894 million of SOV languages. Even when taking statistical error into account (+/- 25%) SVO still outnumbers SOV by far in terms of number of speakers.

Strikingly, this predominance of SVO patterns is mainly due to the fact that the three most widespread languages: Chinese, English and Spanish are SVO languages. Perhaps Chinese has also been subject to pressures from L2 learners? In the case of Modern and Old Chinese, Xu (2006) argues that in earlier periods, Mandarin was a typologically 'mixed language' because it oscillated between verb-object (VO) and object-verb (OV) word orders. However, in a text-count study of written and spoken Modern Mandarin, Sun (1996) found that 90% of the syntactic objects followed the verb (VO), whereas OV with grammaticalized verb-constructions marking agent/patient dependencies only occur in 10% of the sentences, pointing to SVO as the dominant word order. When these observations are combined with our analysis of Classical Latin and Germanic, we may speculate that production pressures from L2 learners can push OV languages with complex systems of solving *who-did-what-to-whom* ambiguities toward a fixed SVO word order with little or no additional case marking. This is consistent with Lupyan and Dale's (2010) statistical analyses indicating that languages with large numbers of users tend to have highly simplified systems of morphology and case.

A second interesting fact about the general trade-off between case marking systems and word order systems was partly anticipated by Greenberg (1966) in his *Universal 41*: 'If in a language the verb follows both the nominal subject and the nominal object as the dominant order, the language almost always has a case system.' Or, expressed in WALS terms: If a language can

Figure 4.5   WALS distribution of word order patterns mapped onto the number of speakers
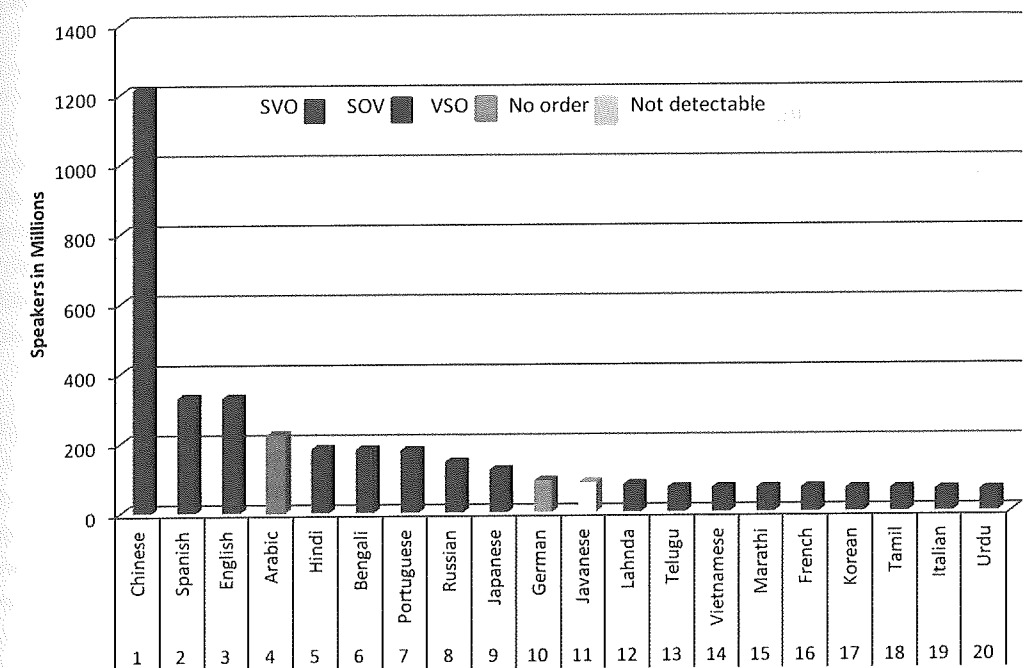

Figure 4.6   Mapping the number of SOV and SVO languages onto the number of case categories attested in the World Atlas of Language Structures (Haspelmath et al. 2005)

The category "borderline case marking" was left out because Iggesen (2011) argued that these are not genuine case markers. For the 4 languages with OSV order in the WALS there is no information on case marking available.
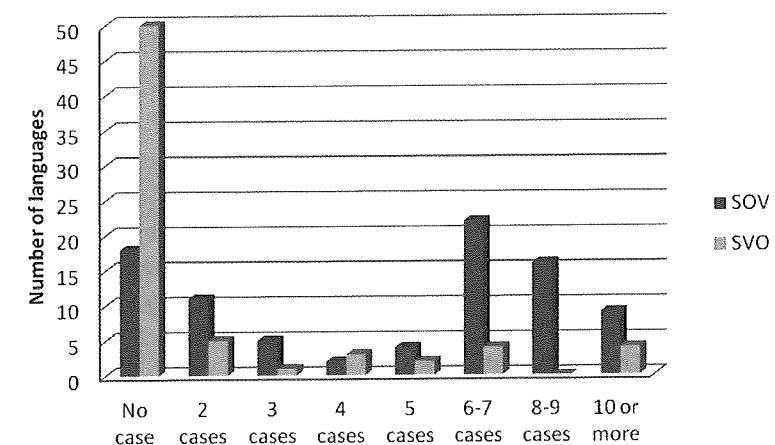
be categorized as either SOV or OSV type, than it should display case marking. This assumption can be tested by crossing the features (49A) 'Number of Cases' and 'Order of Subject, Object and Verb' of the WALS, yielding the distribution shown in Figure 4.6.

There are two interesting trends to be observed in this crossover of word order and case marking. 1) Indeed, there are 98 SOV type languages for which case marking information is available, 70 of these languages (80%) confirm Greenberg's universal, because they display 2 or more cases, whereas only 18 of these languages (20%) do not have case marking at all (contradicting Greenberg's universal). Moreover, especially abundant case marking systems of more than 6 cases seem to be associated with SOV word order (47 languages, 53% of the total). 2) The opposite pattern is true for the 69 SVO languages, 50 of these (72%) are associated with the category 'no morphological case marking', versus 19 (28%) with more or less abundant case marking.

Summarizing these statistical trends, we observe that the association of SOV word order with case marking and the dissociation of SVO languages with case marking are general trends that hold for a wide sample of languages across the world. This is in line with another quantitative study by Bentz and Winter (in press) who aim to show that the impact of L2 learning constraints on case marking paradigms is not restricted to the Romance and Germanic branches of Indo-European languages but rather holds for a wider variety of languages in different families and areas.

6. Conclusion

In this chapter we have argued that the pattern of SVO with no case marking in Modern Romance and Modern Germanic languages has been preceded by stages in proto-languages such as Latin, Proto-Germanic and Proto-Indo-European that still maintained flexible SOV systems with fully fledged case marking paradigms. These nominal marking strategies gradually eroded away because of language internal and external factors such as production constraints of L2 learners, up to a point where the case marking strategy could no longer be maintained. At this point, an alternative strategy—the encoding of who-did-what-to-whom information by a fixed SVO word order—started to dominate sentence structures.

Because there seems to be a trade-off between strict SVO word order without case marking and flexible OV word order with additional morphological markers (Greenberg 1966, Universal 41), it is an interesting fact that both Romance and Germanic languages 'chose' the first strategy. In this paper, we have suggested that this change may be an example of how language adapts to the human brain. In particular, the difficulties in determining the relevant dependency relationships and generating the appropriate sequence of case-marked words would make L2 Latin learners and non-native Middle English speakers prone to errors. L2 production pressures may furthermore have played a role in the similar shift from a relatively flexible word order to fixed SVO in Chinese, which has historically recruited a large number of L2 speakers. The degree to which similar pressures have played a role in the development of a wider sample of

languages from different language families and areas needs to be answered in future research. A first promising attempt has shown that the number of L2 speakers in relation to the L1 speakers of a population can be used as a predictor for the number of case marking paradigms in statistical analysis (Bentz and Winter, in press).

Besides these purely quantitative arguments, it should be noted that pressures from L2 speakers' learning constraints may not only provide an explanation for case loss but may also hint at an explanation for case evolution. In connection to the distinction between exoteric and endoteric languages (Wray and Grace 2007), it seems reasonable to speculate that more closed societies and their languages should be more likely to develop opaque morphological forms such as case markers due to L1 learning constraints, whereas open societies would tend to lose these forms as a result of a shift in the speaker population. Such tendencies would change the cognitive 'niche' within which language is adapting (Lupyan and Dale 2010). This additional explanatory power could be an advantage over traditional accounts of case loss, which do not answer the question of why case marking came into being in the first place.

From a more general point of view, our analyses suggest that historical language change can be used as a source of data for understanding the kind of constraints that may have shaped linguistic adaptation over historical and perhaps even evolutionary timescales. From this point of view, the distinction between language evolution as a biological adaptation of the human language competence on the on hand, and historical language change as a side effect of language performance on the other hand, is misleading in the sense that it disguises the complex diachronic interplay between human learning constraints and language structures.

REFERENCES

Admoni, V. G. 1990. *Historische Syntax des Deutschen*. Tübingen: Niemeyer.

Baugh, A. C., and T. Cable. 2006. *A History of the English Language*. London: Routledge.

Behaghel, O. 1923. *Deutsche Syntax: Eine Geschichtliche Darstellung,* Band I, *Nomen, Pronomen*. Heidelberg: Winter.

Bentz, C., and B. Winter. in press. Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change.*

Braune, W., I. Reiffenstein, and R. Schrodt. 2004. *Althochdeutsche Grammatik* (15. Aufl). Tübingen: Niemeyer.

Campbell, A. 2003. *Old English Grammar* (Repr.). Oxford: Clarendon Press.

Christiansen, M. H. in press. Language has evolved to depend on multiple-cue integration. In *The Evolutionary Emergence of Human Language,* eds. R. Botha, and M. Everaert. Oxford: Oxford University Press.

Christiansen, M. H., and N. Chater. 2008. Language as shaped by the brain. *Behavioral and Brain Sciences* 31:489–509.

Dal, I. 1962. *Kurze deutsche Syntax. Auf Historischer Grundlage* (2., verb. Aufl). Tübingen: Niemeyer.

Ebert, R. P., O. Reichmann, and K.-P. Wegera. 1993. *Frühneuhochdeutsche Grammatik*. Tübingen: Niemeyer.

Ebert, R. P. 1978. *Historische Syntax des Deutschen*. Stuttgart: Metzler.

Greenberg, J. H., ed. 1966. *Universals of Language: Report of a Conference held at Dobbs Ferry, New York, April 13–15, 1961* (2nd Ed.). Cambridge, MA: MIT Press.

Herman, J., and R. Wright. 2000. *Vulgar Latin*. University Park, PA: The Pennsylvania State University Press.

Harris, M. 1988. *The Romance Languages*. London: Routledge.

Haspelmath, M., M. S. Dryer, D. Gil, and B. Comrie, eds. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.

Hawkins, J. A. 2004. *Efficiency and Complexity in Grammars*. New York: Oxford University Press.

Heine, B., and T. Kuteva. 2009. *The Genesis of Grammar. A Reconstruction*. Oxford: Oxford University Press.

Hogg, R. M., and R. Lass. 1999. *The Cambridge History of the English Language*. Vol. III: 1476–1776. Cambridge: Cambridge University Press.

Hutterer, C. J. 2002. *Die germanischen Sprachen. Ihre Geschichte in Grundzügen*. (Nachdr. der 4., erg. Aufl.). Wiesbaden: VMA-Verlag (Albus).

Lee, C. 2000. *Linguistica Romanza*. Rome: Carocci.

Lupyan, G., and R. Dale. 2010. Language structure is partly determined by social structure. *PLoS ONE* 5(1), e8559. doi:10.1371/journal.pone. 0008559

Meier-Brügger, M., M. Fritz, and H. Krahe. 2002. *Indogermanische Sprachwissenschaft*. (8., überarb. und erg. Aufl. der früheren Darstellung von Hans Krahe). Berlin: de Gruyter.

Meillet, A. 1934. *Introduction à l'étude Comparative des Langues Indo-européennes*. Paris: Hachette.

Paul, H., T. Klein, I. Schöbler, and H.-P. Prell. 2007. *Mittelhochdeutsche Grammatik* (25. Aufl.). Tübingen: Niemeyer.

Pinkster, H. 1991. Evidence for SVO in Latin? In *Latin and the Romance Languages in the Early Middle Ages*, ed. R. Wright, 69–82. London: Routledge.

Pintzuk, S. 1999. *Phrase Structures in Competition: Variation and Change in Old English Word Order*. New York: Garland.

Polinksy, M., and E. van Everbroeck. 2003. Development of gender classifications: Modeling the historical change from Latin to French. *Language* 79:356–390.

Slobin, D. I., and T. G. Bever. 1982. Children use canonical sentence schemas. A crosslinguistic study of word order and inflections. *Cognition*, 12:229–265.

Speyer, A. 2007. *Germanische Sprachen. Ein historischer Vergleich*. Göttingen: Vandenhoeck und Ruprecht.

Sun, C. 1996. *Word-order Change and Grammaticalization in the History of Chinese*. Stanford, CA: Stanford University Press.

Wray, A., and G. W. Grace. 2007. The consequences of talking to strangers. Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua* 117:543–578.

Xu, D. 2006. *Typological Change in Chinese Syntax*. Oxford: Oxford University Press.