Harald Clahsen

Contributions of linguistic typology to psycholinguistics

Suggested citation referring to the original publication:
Linguistic Typology 20(3) (2016), pp. 599–614
DOI http://dx.doi.org/10.1515/lingty-2016-0031
ISSN (print) 1430-0532
ISSN (online) 1613-415X

Harald Clahsen

Contributions of linguistic typology to psycholinguistics

DOI 10.1515/lingty-2016-0031
Received August 30, 2016; revised September 25, 2016

Abstract: This article first outlines different ways of how psycholinguists have dealt with linguistic diversity and illustrates these approaches with three familiar cases from research on language processing, language acquisition, and language disorders. The second part focuses on the role of morphology and morphological variability across languages for psycholinguistic research. The specific phenomena to be examined are to do with stem-formation morphology and inflectional classes; they illustrate how experimental research that is informed by linguistic typology can lead to new insights.

1 Introduction

Linguistic typology investigates the universality and variability of the different subsystems of language (from phonetics to pragmatics) across different languages and seeks to chart and understand the nature of linguistic diversity and its limits based on a wide variety of crosslinguistic data. Clearly, this research tradition has accumulated an enormous amount of knowledge and expertise on the world's languages, which informs all the core disciplines of linguistics. But what is the use of linguistic typology for the hyphenated disciplines, the ones in which linguists collaborate with researchers from other fields? Here, I will address this question for psycholinguistics, which investigates (i) the acquisition of language, (ii) language processing in production and comprehension, and (iii) language disorders in children and adults.

Harald Clahsen, Potsdam Research Institute for Multilingualism, Universität Potsdam, Campus Golm, Karl-Liebknecht-Str. 24–25, 14476 Potsdam, Germany,
E-mail: harald.clahsen@uni-potsdam.de

2 Dealing with linguistic diversity in psycholinguistic research

2.1 Does linguistic diversity matter?

One possible viewpoint would be that even though knowledge about linguistic diversity may be useful, it is not essential for understanding the nature of language acquisition, disorders, and processing. After all, psycholinguistic and neurocognitive models specify the mechanisms and mind/brain systems involved in language acquisition, processing, and disorders, and these fundamental mechanisms should be examinable from any human language. Chomsky (1995) made a similar point with respect to the use of linguistic typology for theoretical linguistics. He noted that "the apparent richness and diversity of linguistic phenomena is illusory and epiphenomenal, the result of interaction of fixed principles under slightly varying conditions" (Chomsky 1995: 8), and that the similarities and differences between particular languages do not reflect any deep properties of the language faculty. Although few psycholinguists would subscribe to this viewpoint, there are different approaches in psycholinguistics of dealing with linguistic diversity. Some have ignored or downplayed crosslinguistic differences, whereas others have designed research programmes that explicitly include comparisons of typologically different languages. Consider, for illustration, three well-known cases: the past tense debate (Section 2.2), canonical sentence schemas (Section 2.3), and agrammatism (Section 2.4).

2.2 The past tense debate

My first case is the so-called "past tense debate" (e.g., McClelland & Patterson 2002; Pinker & Ullman 2002), a good example to illustrate the unfortunate focus on English in psycholinguistic research. The past tense debate addresses a core issue of language processing research, namely the role of morphological structure for the mental representation and processing of complex word forms. Are inflected word forms computed from their morphological component parts or are they retrieved and are they stored in the same way as uninflected word forms, through orthographic, phonological, and semantic codes, without any direct representation of their morphological structure? Evidence from a large number of psycholinguistic studies on the English past tense has shown differences between regular *-ed* forms such as *walked* and irregular forms such as *sang* or

went in data from child language acquisition, adult language processing, brain-imaging and event-related potentials, and patients with language disorders indicating that a simple yes/no answer to the above question seems inappropriate. Instead, these contrasts have been taken as support for dual-morphology models, e.g., the "words-and-rules" account (Pinker 1999) in which regularly inflected forms are thought to be represented and processed in terms of their morphological structure (e.g., [walk]+[-ed]) whereas irregularly inflected forms are represented in a lexical memory system of morphologically unstructured units.

The past tense debate came about when single-mechanism models were proposed that did not encode any kind of morphological structure but were still capable of modelling differences between regular and irregular past tense forms of English observed in psycholinguistic data. For example, children's overgeneralization errors are with regular (rather than irregular) inflectional forms (e.g., **bringed* instead of *brought*, but not **truck* instead of *tricked*). Single-mechanism models (see Rumelhart & McClelland (1986) and much subsequent work) were able to mimic this contrast in computational models that had no knowledge of any kind of morphology or morphological rules. Note, however, that these models exploited a number of unfortunate confounding variables of English past tense inflection, firstly, that there are many more verbs that take *-ed* forms than irregular past tense forms (95% vs. 5%; see Marcus et al. (1995)), and secondly, that regular forms are pure concatenative forms whereas irregular forms typically involve phonological changes and no overt affix. These factors allow single-mechanism models to mimic human behaviour on inflected words without encoding any kind of morphological information. For example, Rumelhart & McClelland's (1986) pattern-association model appears to exhibit productive generalizations of *-ed* past tense formation, simply because *-ed* forms represent the most frequent pattern. Likewise, selective impairments of *-ed* past tense formation, as reported for patients with agrammatic aphasia (e.g., Ullman et al. 2005) for example, were supposed to be replicable in a single-mechanism model through disrupted phonological links (Joanisse & Seidenberg 1999), again taking advantage of an idiosyncratic contrast between regular and irregular past tense forms in English, namely that *-ed* forms are phonologically more faithful than irregulars (e.g., *think* – *thought*). Grand claims were made as a result of these models, for example, that language performance can be understood without reference to morphological structure (and perhaps more generally – grammar) and that language knowledge and use can better be represented through associatively linked orthographic, phonological, and semantic codes. A less sensational but perhaps more appropriate conclusion from the past tense debate would have been that, due to its various confounding

factors, the English past tense is perhaps not an ideal case for testing psycholinguistic models of morphology, and that crosslinguistic studies and insights from typologically different languages would have been useful in this case – and would have taken away a lot of unnecessary steam from the heated past tense debate.

2.3 The canonical sentence schema

My second case comes from language acquisition research, which – thanks to the project "The cross-linguistic study of language acquisition" initiated by Dan Slobin – has gone well beyond the study of English and has since the late 1970s accumulated a wealth of knowledge on child first language acquisition across a wide range of typologically different languages; see Berman (2014) for a recent review. Here, I will rely on results from this research project to illustrate how crosslinguistic comparisons without the necessary linguistic sophistication may lead to spurious claims and the wrong conclusions. Consider the notion of a "canonical sentence schema" (Slobin & Bever 1982) as an example. The canonical sentence schema is defined as the most common sentence template in a given language for a simple active affirmative declarative sentence involving a transitive verb and the related subject and object nouns (Slobin & Bever 1982: 253). Assuming that "the most common patterns are learned first" (Townsend & Bever 2001: 372), Slobin & Bever (1982) posit that by two to three years of age, typically-developing children will have learnt the canonical sentence schema for their particular language, which they will use for both production and comprehension. They point to English and Turkish child language to support their account, with English-speaking children identifying the S(ubject)-V(erb)-O(bject) and Turkish children the SOV schema of their respective languages early on.

Note, however, that matters for other languages are a bit more complicated. By Slobin & Bever's (1982) definition, the canonical sentence schema of German is S-V-O, as in English. However, as shown in studies of German child language (Clahsen 1986 and much subsequent work), two-to-three-year-old German children predominantly produce sentences with (S)OV order along with sentences in (S)VO order, with the crucial difference that – corresponding to the structure of main clauses in the adult language – the "V(erbs)" in the children's (S)OV sentences are typically non-finite and the ones in the (S)VO sentences finite. The crucial point here is that children are not just scanning the relative surface ordering of subjects, verbs, and objects, but that – if necessary – they are also sensitive to morphological distinctions early on, in the case of German, the

different ordering patterns of finite and non-finite verbs. While children seem to detect the OV order of German early on, it takes time for them to learn the morphology necessary to productively form finite verbs. This is why at early stages only few sentences contain a finite verb, a straightforward developmental path; see Clahsen & Penke (1992). In any case, linguistically oversimplified notions (e.g., "canonical sentence schema") may lead to the wrong conclusions, not only in linguistic typology but also in crosslinguistic research on language acquisition.

2.4 Agrammatism

My third case comes from research on language disorders, and specifically on "agrammatism", the main symptom of Broca's aphasia. Agrammatism is commonly defined as "speech production in which many function words and inflectional endings are omitted" (Field 2004: 7). This definition does indeed apply to languages such as English in which morpheme omission yields legal word forms. While omitting *-ed* from *walked*, for example, produces a well-formed word in the English language, this does not hold for many other languages in which the omission of inflectional morphemes would lead to phonotactically illegal or even unpronounceable strings. Speakers of such languages with Broca's aphasia do indeed refrain from omitting inflectional morphemes but instead produce unmarked forms such as infinitives, or substitution errors, as shown, for example, by Grodzinsky (1990) for agrammatism in Italian, Russian, and Hebrew. To take an example from Grodzinsky (1990), Italian-speaking patients with agrammatism produce gender-agreement errors such as **Cappuccetto Rossa* rather than bare roots (*Cappuccett- Ross-*) instead of the correctly inflected forms (*Cappuccetto Rosso* 'Little Red Riding Hood'). In any case, a defining property of agrammatic aphasia that even went into clinical diagnosis and practice turned out not to hold crosslinguistically.

3 Insights from linguistic typology for psycholinguistic research on morphology

3.1 Stems and inflectional classes

Psycholinguistic research on morphologically complex words is typically confined to a restricted set of largely morpheme-based phenomena (e.g., the past

tense) that are analysed using binary distinctions (e.g., "regular" vs. "irregular") and a limited number of morphological types, most commonly "word" and "affix". This approach works relatively well for inflection, derivation, and compounding in English and other Germanic languages. If it is applied to other languages, however, the wrong conclusions may be drawn while missing out relevant generalizations. Here I will use stem-formation morphology and inflectional classes as an example.

Consider a number of psycholinguistic studies on verb inflection that more or less directly applied the familiar regular/irregular distinction from English to Romance and Slavic languages. De Diego Balaguer et al. (2005) reported that patients with agrammatic Broca's aphasia were more accurate with regular than irregular verb morphology in both Catalan and Spanish, a pattern of performance they claim is inconsistent with the view that regular inflection is specifically impaired in Broca's aphasia (e.g., Ullman et al. 2005). Orsolini et al. (1998) reported more overapplication errors for irregular than for regular verbs forms in Italian children's production of past tense and past participle formation, which they interpreted as evidence against the view that children's overgeneralization errors largely rely on regular (and only rarely on irregular) morphological processes (Marcus et al. 1992). Gor & Cook (2010) investigated what they called regular, semi-regular, and irregular types of Russian verb forms in an auditory lexical decision task and obtained parallel effects across the three conditions, which they took to be "inconsistent with either the dual-system or the single-system approach to morphological processing" (Gor & Cook 2010: 88) which would have predicted stronger effects for regular than for irregular forms. The relevant typological insight here is that – in addition to word formation (viz. derivation and compounding) and inflection – many languages have stem-formation morphology. Stem formation makes use of different kinds of exponent including thematic suffixes, stem formatives, inflectional class markers, and others; see Nichols (2016) for a recent review. Good examples of stem-formation morphology can be found in Romance and Slavic languages.

The problem with the three experimental studies mentioned above is that they confuse the properties of stem-formation morphology with those of inflectional processes, even though in neither case is there a direct mapping between the two sets of properties in these languages. Hence, directly applying the notions of regular vs. irregular inflection from English to the phenomena tested is misguided. De Diego Balaguer et al. (2005), for example, found that inflected forms with marked stems such as *salgo* '(I) leave' from *salir* 'to leave' elicited more errors than verbs with unmarked stems such as *miraba* 'looked' from *mirar* 'to look'. This result simply indicates that patients with aphasia have more difficulty retrieving marked than unmarked stems, a contrast that is familiar

from Spanish child language (Clahsen et al. 2002) and likely to occur in other "challenged" populations. Crucially, this contrast has nothing much to do with regular vs. irregular inflection, but reflects contrasts between different stem types that also occur for derivational morphology in Spanish and other languages. Consequently, unlike to what De Diego Balaguer et al. (2005) claim, their findings do not bear on theories that posit a specific deficit in regular inflection for patients with Broca aphasia. Orsolini et al. (1998) expected Italian children to only use first conjugation verb forms in their morphological errors, because only the first conjugation is regular. It is true that in Romance languages only the first conjugation shows unrestricted productivity, but verbs belonging to second or third conjugation classes may still be subject to "regular" inflectional processes in the same way as first conjugation verbs. Thus, second conjugation past tense forms are not "irregular" qua their conjugation class, as Orsolini et al. (1998) wrongly claimed. Say & Clahsen (2001) reanalyzed Orsolini et al.'s child language data and found that almost all of the errors were overgeneralizations of unmarked stems and regular affixation rules, indicating that their data are in fact compatible with the dual-morphology account. Likewise, in Gor & Cook's (2010) study, the inflectional processes involved in all their word pairs were parallel, whereas the nature of the stems differed, with the so-called regular condition involving the same stems in both items (*rabotaju – rabotat'* 'to work'), the "semi-regular" condition containing one phonologically predictable alternation (*xožu – xodit'* 'to walk'), and the "irregular" condition containing one unpredictable alternation (*kladu – klast'* 'to put'). Hence, their experiment did not test inflectional processes but rather manipulated different types of stem, which apparently had no or little effect on participants' response times.

That the distinction between stem formation versus word formation and inflection matters can be seen from the kinds of morphological errors children produce in their speech. Consider mixed errors in Spanish and Greek child language. Three-to-four-year old children have been reported to produce errors such as **pusí* [I.put.PST] (correct: *pus-e*) in Spanish or **eplise* instead of *epline* [washed.3SG] in Greek (Clahsen et al. 2002; Stavrakaki & Clahsen 2009). In such errors, marked (irregular) stems of the corresponding verbs (e.g., *pus-*) are combined with regular inflectional suffixes (instead of irregular ones) indicating that children can manipulate stems and inflectional endings separately. Another type of error produced by Spanish-speaking children is illustrated by examples such as **devolvido* instead of the correct form *devuel-t-o* [return-PTCP-M.SG] in which the child produced the unmarked root (instead of the diphthongized one) and the regular participle suffix *-d-*. Of interest here is the use of theme vowels in these kinds of error. In the above example, the theme vowel supplied

by the child is indeed correct, despite the errors in the root and the suffix; *devolv-* is a second conjugation root for which the required theme vowel is *-i-*. This pattern holds for all other errors of this kind in that there was no single case of a conjugation class error in such cases. For example, errors such as **devolv-a-do*, i.e., a regularized form of *devuel-t-o* with the first conjugation theme vowel *-a-* were nonexistent in the large data set examined by Clahsen et al. (2002). These cases indicate that in children's morphological errors the formation of stems can be dissociated from root-related and inflectional processes.

The English-influenced approach of labelling whole-word forms as either "regular" and "irregular" is unfortunate and inappropriate for many languages. Instead, stem-formation morphology should be considered separately from other derivational and inflectional processes. In addition to words and roots, stems have been argued to represent an independent morpholexical type (e.g., Aronoff 1994; Anderson 1992; Stump 2001). While roots are morphologically unanalyzable and are defined with respect to lemmas, stems may be morphologically complex and represent recurrent parts of a series of inflected and/or derived forms. Stems are widespread – perhaps even universal (Berent et al. 2007) – across human languages. Stem formation can be COMBINATORIAL, yielding, for example, combinations of a root plus a stem-formation affix (e.g., Italian: [[*cant*]-*a*] in *cantare* 'to sing') or ALLOMORPHIC involving, for example, vowel changes (e.g., German: *werfen* [throw.INF] – *warf-* [throw.PST], or even SUPPLETIVE as, for example, in past tense forms of *sein* 'to be' in German (e.g., *war-st* [be.PST-2SG], *war-en* [be.PST-1/3PL]). Some stem-formation processes encode morphosyntactic features such as tense, person, number, etc., others are nonmorphemic without any specific meaning or morphosyntactic feature content defining, for example, conjugational or declensional classes. German verb stems such as *warf-* and *war-* are of the former, Italian verb stems such as [[*cant*]-*a*] of the latter type. Once this distinction is made, we can ask relevant psycholinguistic questions about different kinds of stem-formation processes, such as how productive they are, how they are mentally represented, and how they are processed during production and comprehension. While to date there are only few experimental studies that have investigated stem-formation morphology from this perspective, these studies have produced new and partly surprising insights that are also of interest to the general linguist.

3.2 A case study: Experimental studies on stems and conjugation classes in Portuguese

Consider results from experimental studies on Portuguese (Veríssimo & Clahsen 2009, 2014) that examined how different kinds of conjugational stem – as found

in Romance, Slavic, and many other languages – generalize to novel verbs and how existing verb forms of different conjugational classes are processed during word recognition. One crucial property of morphological processes is that they are productive to different degrees, i.e., generalizable to novel or unusual words under appropriate circumstances. Generalization properties provide a crucial diagnostic for how morphological categories and processes are mentally represented. Morphological productivity has previously mainly been investigated with respect to inflectional affixes such as the English past tense *-ed* (e.g., Bybee & Moder 1983; Prasada & Pinker 1993). Relatively little is known to date about the generalization properties of different types of stem.

To examine the productivity of the different conjugational stems in Portuguese, Veríssimo & Clahsen (2014) performed an elicited production task in which participants were presented with novel verbs in 1st person singular present tense indicative forms (which do not provide any clues to conjugation class) and had to produce infinitives (by necessarily assigning the root to one of the three verbal conjugations). For example: *Quase sempre acuo sozinho. Mas amanhã vou ___ acompanhado*. 'I almost always *acuo* alone. But tomorrow I will ___ with someone.' With the aim of systematically controlling for phonological similarity, novel verb forms were constructed using a computational model, the Minimal Generalization Learner (MGL) (Albright 2002; Albright & Hayes 2003) applied to the Portuguese verb lexicon. Novel verbs were selected for the experiment representing phonological environments that the MGL model identified as constituting particularly reliable contexts for one or more conjugational classes. The results of the elicited production experiment revealed a clear dissociation between the three conjugational stem types. Proportions of second and third conjugation responses were solely determined by phonological similarity, i.e., their corresponding MGL reliability values. In contrast, phonological similarity with the first conjugation did not reliably predict first conjugation responses. Instead, the likelihood of a first conjugation response was predicted by phonological similarity to the second and third conjugations, such that the higher the similarity for these classes, the lower the proportion of participants producing first conjugation stems. Veríssimo & Clahsen (2014) also demonstrated that a purely similarity-based model fails to account for the human participants' performance, in that this kind of model consistently underestimated the proportion of responses with first conjugation and overestimated those with second and third conjugation stems. Instead, these results provide support for a dual-morphology account of conjugational stem representation by which first conjugation stems are taken to be the output of morphological rules with unlimited productivity while second and third conjugation stems comprise sets of items stored in the mental lexicon; see Say & Clahsen (2002) for Italian.

This contrast accounts for their different generalization properties: first conjugation stem formation generalizes by default to all members of a given grammatical category, whereas second and third conjugation stem forms generalize by analogy, depending on a novel word's phonological overlap with existing verbs of these classes.

Psycholinguists also investigate how morphological properties affect the recognition and comprehension of complex words. To gain detailed insight into these processes and their temporal sequencing, time-sensitive experimental measures are required. Lexical priming techniques are particularly suited for this purpose as they are sensitive to the internal structure of a complex word form. In morphological priming tasks, participants are presented with a morphologically complex prime word before a different word form as a target word, e.g., *walked* as a prime for the target *walk*, for the latter of which they perform a lexical (word vs. nonword) decision task. The rationale behind priming experiments is that shorter reaction times to the target word following a related prime word indicate that the target word (e.g., the stem *walk*) was pre-activated during the recognition of the prime word. While morphologically related prime words generally lead to shorter response times on target words than unrelated primes (e.g., *walked – walk* vs. *looked – walk*), a particularly striking finding is a full priming effect, in which case the amount of facilitation on target recognition times after a morphologically related word form is of the same magnitude as for an identity prime (e.g., *walked – walk* vs. *walk – walk*). Full priming is obtained when the same lexical unit is activated by prime and target, which for morphologically complex words is believed to require some kind of morphological analysis prior to memory access. Full priming for pairs such as *walked – walk*, for example, has been supposed to be due to the prime being morphologically decomposed during recognition ([*walk*]-*ed*), thereby isolating the base stem, which then directly facilitates recognition of the target word *walk*, in essence a stem-repetition effect; see Stanners et al. (1979) and much subsequent work.

Veríssimo & Clahsen (2009) examined different kinds of conjugational stems in Portuguese using a morphological priming experiment in which infinitive forms belonging to the first (e.g., *limit-a-r* 'to limit') or the third conjugation (e.g., *resist-i-r* 'to resist') were compared as primes for corresponding root-based present tense indicative forms (e.g., *limit-o, resist-o*). The most important finding from this study was that – despite the fact that the primes and targets in both conditions were parallel in terms of their orthographic, phonological, and semantic relatedness – stem forms of the first and the third conjugation produced different priming effects on root-based target forms: full priming for first conjugation verbs and partial priming for third conjugation verbs. This finding challenges the familiar three-way distinction between first, second, and third

conjugation in Portuguese, Italian, and most probably other Romance languages that posit the same representations for the verbal stems of the different conjugations; see, for example, Villalva (2000) for Portuguese. Instead, the contrast obtained between first and third conjugation stems can be better explained by positing different representations for these stems, root-based morphologically-structured ones for the former, and unstructured stem-based representations for the latter (Say & Clahsen 2002). Assuming that full priming is indicative of repeated activation of the same lexical unit in prime and target, the full-priming effect seen for first conjugation stems indicates that these stems are decomposed down to the root (e.g., [*limit*-[*a*]-]). By contrast, the reduced priming effect obtained for third conjugation verb forms suggests that they have unstructured stem representations (e.g., [*resisti*-]), from which the corresponding root can only be indirectly activated. Taken together, the results from both elicited production and lexical priming experiments provide evidence for a dual morphology account of conjugational stems that distinguishes between rule-based computed stems (e.g., for first conjugation stems in Portuguese) and listed morphologically unstructured stems (e.g., third conjugation stems in Portuguese). At a more general level, these findings can be explained in terms of the general opposition between items that are stored in the lexicon and elements that are derived from lexical items by rule-like operations, a contrast familiar from dual-morphology models (e.g., Pinker 1999: 237). The results reported above indicate that this account can be extended to stem-formation morphology.

3.3 Roots, stems, and *binyanim*: Experimental studies of Semitic morphology

Another topic of psycholinguistic research on morphology for which consultation of linguistic typology may be beneficial concerns the question to what extent mental representation and processing of morphologically complex is universal and to what extent it is shaped by particular morphological properties of individual languages. The role of stem-formation morphology in Indo-European vs. Semitic languages has featured prominently in psycholinguistic studies on this topic. It has been proposed that morphological processing in Semitic languages is substantially different from processing of other languages, due to the unique nonconcatenative morphology of these languages; see Bick et al. (2011) for Hebrew and Boudelaa & Marslen-Wilson (2015) for Arabic. Unlike in Indo-European languages, many complex words in Hebrew or Arabic do indeed involve nonlinear combinations of consonantal roots – typically consisting of three consonants that carry the core lexical meaning – plus patterns

consisting of vowels to encode grammatical features. These properties have been highlighted as the cause for Semitic morphology to be processed differently from other language types. The Semitic language processor has been said to be primarily "morphological" in nature, designed to extract a complex word's root and word pattern structure irrespective of its meaning or surface form, effectively employing full morphological parsing ("down-to-the-root") as the dominant processing mechanism. By contrast, the system for processing complex words in English, for example, is thought to be less purely morphologically driven and instead more affected by nonmorphological factors such as the semantic transparency and the orthographic and phonological surface form of a complex word as a whole (Bick et al. 2011). To take an example from morphological priming studies, semantically opaque and/or phonologically altered forms (e.g., *kept – keep, business – busy*) typically do not produce facilitation effects in (overt) morphological priming experiments in English (e.g., Marslen-Wilson et al. 1994), but in Hebrew and Arabic they do (Frost et al. 2000). According to Bick et al. (2011), this is due to the prevalence of phonological and semantic opaqueness of morphologically complex words in English and other Indo-European languages, unlike in Semitic languages in which straightforward root-and-pattern extraction is possible for almost all words.

It is true that Semitic languages have nonconcatenative morphology. Note, however, that many languages have nonconcatenative exponents, e.g., stem allomorphy in Indo-European languages, and that vice versa, many inflected words in Semitic languages carry concatenative affixes, e.g., subject-verb agreement suffixes. In any case, there is nothing in the surface encoding of Semitic morphology that requires any kind of special processing mechanisms or that would yield a distinct Semitic morphological parser. As regards morphological types, the crucial linguistic insight is that like many other languages, Hebrew and Arabic make use of lexeme-formation processes to derive stems from roots, for example for Hebrew, the so-called *binyanim* for verbs and the *mishkalim* for nouns. For verbs the *binyanim* are obligatory, much like the conjugation classes of Latin and the languages derived thereof; see Aronoff (1994: 124–128). What is different in Hebrew is that the inflection-class marker is abstract, unlike in Latin in which the formation of verbal stems typically involves affixation (as, e.g., in [*labor*]*a*]- 'work').

With this insight in mind, we can ask sensible psycholinguistic questions for the processing of Semitic morphology. Regarding verbal stem-formation processes in Hebrew, for example, we may examine whether the Semitic mental lexicon is indeed fully decompositional, with inflected or derived words normally segmented down to the root, or whether this is perhaps restricted to particular inflectional classes, similarly to what Say & Clahsen (2002) and

Veríssimo & Clahsen (2009) found for Italian and Portuguese, respectively. A recent study (Farhy et al. 2016) addressed this question for inflected verb forms of Hebrew that belong to the Pa‘al or the Pi‘el *binyan* using the morphological priming technique. While both *binyanim* are very common in the Hebrew verb lexicon, the Pa‘al *binyan* essentially represents a closed class of items, whereas Pi‘el readily welcomes new verbs that enter the language. Farhy et al. (2016) tested two inflected forms as primes (infinitives, 1st person singular past forms) with 3rd person singular past forms as targets. Primes were presented either in Pa‘al or Pi‘el and targets in the Hitpa‘el *binyan*. For example: נישקתי/נשקתי → התנשק (*NaShaKTI*/*NIShaKTI* ‘I kissed’→ *HiTNaSheK* ‘he kissed (reciprocal)’. The target word forms for the two prime conditions (infinitives, 1SG past forms) as well as the Pa‘al and Pi‘el primes within each condition were closely matched with respect to length, frequency, and with respect to their semantic relatedness to Hitpa‘el verbs. The main finding from this study was that despite the fact that both Pa‘al and Pi‘el forms shared the same verbal root with the targets, only primes belonging to the Pi‘el class elicited a reliable root priming effect. That is, prior presentation of a Pi‘el form significantly facilitated recognition of a Hitpa‘el form of the same root, whereas this was not the case for Pa‘al forms. This dissociation between verbal classes was obtained for both infinitive and 1SG past forms. The root-priming effect for Pi‘el forms indicates that these forms are indeed parsed down to the root. The recognition of the Pa‘al forms tested, however, seems to depend on access to full stems from which roots are not automatically extracted, hence the lack of any significant root-priming effect for Pa‘al forms. The contrast obtained between the Pa‘al and Pi‘el classes provides support for the dual-morphology account according to which closed-class forms (in the present case, Pa‘al stems) are stored as wholes while productive forms are computed by rule-like mechanisms, in the present case through a stem-formation rule that generates Pi‘el forms from consonantal roots. At a more general level, Farhy et al.’s results suggest that there is no “Hebrew brain” or “English brain” (Bick et al. 2011: 2280) but that in both languages, the processing system for morphologically complex words makes use of the same mechanisms.

4 Conclusion

The main point from the studies reported above is that while languages differ with respect to the morphological types they typically rely on (e.g., roots, stems, words, affixes), the processing system is flexible enough to handle this kind of variability. Instead, what really seems to matter to the mental representation

and processing of morphologically complex forms is the (probably universal) distinction between forms that constitute units of lexical storage and forms that are computed from smaller units by grammatical rules or equivalent operations.

I conclude that insights from linguistic typology are indeed beneficial for psycholinguistic research in that they may help (i) to prevent illusions, by discovering that supposedly general notions and accounts developed for one language may not directly translate to other languages, (ii) to enrich the scope of research, by pointing to phenomena that are potentially revealing but unavailable from the most commonly studied languages, and (iii) to lead to new insights, by disentangling universal from language-particular properties in acquisition, processing, and disorders.

Acknowledgements: Supported by an Alexander-von-Humboldt-Professorship. I wish to thank the members of the Potsdam Research Institute for Multilingualism (PRIM), particularly João Veríssimo, Yael Farhy, and Claudia Felser for helpful comments on the present article.

Abbreviations: 1/3 = 1st/3rd person; INF = infinitive; M = masculine; PST = past; PTCP = participle; SG = singular.

References

Albright, Adam. 2002. Islands of reliability for regular morphology: Evidence from Italian. *Language* 78. 684–709.

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90. 119–161.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.

Berent, Iris, Vered Vaknin & Gary F. Marcus. 2007. Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition* 104. 254–286.

Berman, Ruth A. 2014. Cross-linguistic comparisons in child language research. *Journal of Child Language* 41(S1). 26–37.

Bick, Atira S., Gadi Goelman & Ram Frost. 2011. Hebrew brain vs. English brain: Language modulates the way it is processed. *Journal of Cognitive Neuroscience* 23. 2280–2290.

Boudelaa, Sami & William D. Marslen-Wilson. 2015. Structure, form, and meaning in the mental lexicon: Evidence from Arabic. *Language, Cognition and Neuroscience* 30. 955–992.

Bybee, Joan L. & Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59. 251–270.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
Clahsen, Harald. 1986. Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics* 24. 79–122.
Clahsen, Harald, Fraibet Aveledo & Iggy Roca. 2002. The development of regular and irregular verb inflection in Spanish child language. *Journal of Child Language* 29. 591–622.
Clahsen, Harald & Martina Penke. 1992. The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-Corpus. In Jürgen M. Meisel (ed.), *The acquisition of verb placement: Functional categories and V2 phenomena in language acquisition*, 181–223. Dordrecht: Kluwer.
De Diego Balaguer, Ruth, Nuria Sebastián-Gallés, Begoña Díaz & Antoni Rodríguez-Fornells. 2005. Morphological processing in early bilinguals: An ERP study of regular and irregular verb processing. *Cognitive Brain Research* 25. 312–327.
Farhy, Yael, João Veríssimo & Harald Clahsen. 2016. Universal and particular in morphological processing: Evidence from Hebrew. Unpublished manuscript, Universität Potsdam.
Field, John. 2004. *Psycholinguistics: The key concepts*. Abingdon: Routledge.
Frost, Ram, Avital Deutsch, Orna Gilboa, Michal Tannenbaum & William D. Marslen-Wilson. 2000. Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition* 28. 1277–1288.
Gor, Kira & Svetlana Cook. 2010. Nonnative processing of verbal morphology: In search of regularity. *Language Learning* 60. 88–126.
Grodzinsky, Yosef. 1990. *Theoretical perspectives on language deficits*. Cambridge, MA: MIT Press.
Joanisse, Marc F. & Mark S. Seidenberg. 1999. Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences* 96. 7592–7597.
Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese & Steven Pinker. 1995. German inflection: The exception that proves the rule. *Cognitive Psychology* 29. 189–256.
Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, Fei Xu & Harald Clahsen. 1992. *Overregularization in language acquisition* (Monographs of the Society for Research in Child Development 228). Chicago: University of Chicago Press.
Marslen-Wilson, William D., Lorraine K. Tyler, Rachelle Waksler & Lianne Older. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review* 101. 3–33.
McClelland, James L. & Karalyn Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6. 465–472.
Nichols, Johanna. 2016. Morphology in linguistic typology. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology*. Cambridge: Cambridge University Press.
Orsolini, Margherira, Rachele Fanari & Hugo Bowles. 1998. Acquiring regular and irregular inflection in a language with verb classes. *Language and Cognitive Processes* 13. 425–464.
Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York, NY: Basic Books.
Pinker, Steven & Michael T. Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Sciences* 6. 456–463.
Prasada, Sandeep & Steven Pinker. 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8. 1–56.

Rumelhart, David E. & James L. McClelland. 1986. On learning the past tenses of English verbs. In David E. Rumelhart, James L. McClelland & The PDP Research Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. 2, 216–271. Cambridge, MA: MIT Press.

Say, Tessa & Harald Clahsen. 2002. Words, rules and stems in the Italian mental lexicon. In Sieb Nooteboom, Fred Weerman & Frank Wijnen (eds.), *Storage and computation in the language faculty*, 93–129. Dordrecht: Kluwer.

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

Stanners, Robert F., James J. Neiser, William P. Hernon & Roger Hall. 1979. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18. 399–412.

Stavrakaki, Stavroula & Harald Clahsen. 2009. The perfective past tense in Greek child language. *Journal of Child Language* 36. 113–142.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Townsend, David J. & Thomas G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Ullman, Michael T., Roumyana Pancheva, Tracy Love, Eiling Yee, David Swinney & Gregory Hickok. 2005. Neural correlates of lexicon and grammar: Evidence from the production, reading, and judgment of inflection in aphasia. *Brain and Language* 93. 185–238.

Veríssimo, João & Harald Clahsen. 2009. Morphological priming by itself: A study of Portuguese conjugations. *Cognition* 112. 187–194.

Veríssimo, João & Harald Clahsen. 2014. Variables and similarity in linguistic generalization: Evidence from inflectional classes in Portuguese. *Journal of Memory and Language* 76. 61–79.

Villalva, Alina. 2000. *Estruturas morfológicas: Unidades e hierarquias nas palavras do português*. Lisboa: Fundação Calouste Gulbenkian & Fundação para a Ciência e a Tecnologia.