Compounds and Productivity in Advanced L2 German Writing: A Constructional Approach

Amir Zeldes
Georgetown University

Abstract

The frequent formation of complex, hierarchically structured compounds is a striking property of German grammar to non-natives. This article asks how compounding works in second language (L2) German grammar, by exploring data from the error-annotated Falko corpus of native and advanced non-native German writing. Beyond differences in overall frequency and productivity of L2 compounding, I use a constructional approach based on compound paraphrases and partially filled prototypes to analyze differences between first language (L1) and L2 usage, as well as identify frequent error types. The results reveal that productivity as evidenced by rare items in L2 output is a key factor in the native-like acquisition of compounding, and that proficiency as assessed by a C-Test correlates better with more complex productivity measures than with raw vocabulary size. The data also show significant differences in compounding frequency based on learner native language and some possible phonetic explanations for morphological errors at the boundary between compound heads and modifiers. Semantic errors are overall rare but in many cases attributable to transfer effects, even from constructions that are not compounds in the underlying L1, or indeed from languages low in compounds. This suggests that both abstract and partially lexicalized compounding constructions are learned and errors can affect either of these at the lexical level.

Introduction

The productivity, complexity, and sheer length of nominal compounds in German is a striking characteristic of German grammar in usage (see Schlücker 2012). Although compounding is typologically a frequent phenomenon, German is considered unusually 'compound-happy' within the Indo-European family, and is even extreme in its frequency by Germanic standards (Neef 2009:386). The examples below demonstrate native usage, taken from the Falko Essay control corpus (see Reznicek et al. 2012 and details below).[1]

(1) *noch funktioniert es in aufgeblasenen Beamte**n**verwaltung**s**apparaten*
still functions it in inflated civil-servant.administration.apparatuses
'it is still working in inflated civil servant administration apparatuses'

[1] I use periods to separate compound constituents in glosses. Linking elements are marked with underlines in the German.

(2) *wenn Lehre und Forschung in eine "Drittmittelbesorgung[**s**]notwendigkeit" hineingerät*
when teaching and research in a third.resource.acquisition.necessity into-falls
'when teaching and research falls into an "external funding acquisition necessity"'

The complexity of such compounds lies not only in superficial factors such as orthography without spaces, but also:

- Frequent bracketing ambiguity ([civil servant [administration apparatuses]] or [[civil servant administration] apparatuses]).
- The only partially predictable morphology of linking elements, marked above by underlines (see Nübling & Szczepaniak 2013).
- Different parts-of-speech: *drittmittel* 'third resource (=third-party)': A+N; *Verwaltungsapparaten* 'administration apparatuses': N+N.
- Different constructions licensing semantic readings of compounds, including argument structure (in *besorgung* + *notwendigkeit*, the 'being_necessary' frame invokes a semantic 'requirement' argument in Frame Semantics).[2]

These factors may seem intuitive, since speakers of a language typically negotiate nesting, argument structure, parts-of-speech and morphology outside of compounding too, but in fact even native speakers show variation in compound realization and violations of normative grammar (e.g. missing linking *[s]* in the last example). For non-natives the challenges are naturally more substantial.

Although the prevalence and difficulty of German compounding has been commented on often, there are comparatively few theoretical studies on how compounding is used and what constructions are acquired in L2 German. The importance of this topic is augmented by the fact that compounding is taught and also acquired early by L2 learners of German, and forms a major strategy for expanding vocabulary early on (Heidermann 1997:58-60, Fandrych & Tallowitz 2008:202). A substantial amount of previous work has focused on optimizing dictionaries and grammars for (beginning) learner use (see Rak 1990, Fandrych & Thurmair 1994, Barz 1995, Trynkler 2010), but comparatively little has been said about how more proficient learners actually use compounds once they are able to use them productively (see Zeldes 2013), or about how their use of compounds may or may not differ from their native counterparts in practice.

This chapter focuses directly on advanced L2 German learners' use of compounds. In particular, I will be interested in the following theoretical question: What is it that a learner acquires when they learn to form (novel) German compounds? How does it differ from native speakers' knowledge? Some options that will be considered and contrasted are:

- Recursive abstract rules of the type $[\text{N-N}]_{\text{N}}$
- More specific abstract constructions of the type $[\text{N-[V]}er_{\text{N}}]_{\text{N}}$ (argument structure compounds of the pattern 'noun-verb-er', e.g. *Taxifahrer* 'taxi driver')

[2] Cf. FrameNet, https://framenet.icsi.berkeley.edu/. See Gawron (2011) for an overview.

- Semi lexicalized patterns forming 'positional families' (see de Jong et al. 2002) where multiple heads ([N-*Leiter*$_N$]$_N$ 'N-manager') or modifiers ([*Profi*-N]$_N$ 'pro-N, professional-N') form an extensible, prototype-centric series of constructions

In this chapter I will show that while all of these strategies are in evidence in learner data, a great deal of what distinguishes L1 and L2 compound production can be expressed in terms of partially lexicalized constructions, rather than grounded in general rules of the N+N=N type. A further, independent question that will be addressed is whether the ability or tendency to create novel compounds is related to L1 background or to learner proficiency. If proficiency, in particular emerges as a meaningful variable, I will want to ask whether productive compounding is more native-like in more proficient learner data. I will also show that the data can be analyzed in terms of a hierarchy of progressively abstract constructions, which I will describe within a constructional model, taking elements from syntactic and morphological research in Construction Grammar (Goldberg 2006, Booij 2010).

Theoretical Underpinnings

A discussion of tendencies in L2 compounds necessitates a usage-based approach. Especially for advanced learners, who are more likely to produce compounds productively, differences to L1 behavior are not categorical, beyond clear morphological errors. Rather, they produce forms that diverge from expected target norms for expressing certain functions (Lüdeling & Walter 2010:319), overuse or underuse otherwise appropriate items (especially over-preferring the most frequent member of a category, see Ringbom 1998, Ellis 2012, Ellis et al. 2014) or produce combinations that are unlikely, but not strictly unacceptable (Nesselhauf 2003). If we are interested in differences between L1 and L2, we must examine usage frequencies, usage contexts and systematic patterns within these. As a result, I argue we must adopt a usage-based approach, which grants the systematicities we observe a status within the grammar (or each learner's, and each L1 or L2's interlanguage grammar, following de Angelis & Selinker 2001).

The term Construction Grammar unites a wide range of approaches (Croft 2001, Fillmore 1988, Goldberg 2006) which have in common at least:

- Assuming information about usage and frequencies is stored in the mental lexicon (e.g. 'entrenchment', cf. Langacker 1987, Bybee 2006)
- Allowing for the existence of 'constructions', arbitrary form-meaning pairings which can be fully lexicalized, only partially lexically specified or fully abstract
- Meaning contributions in compositional cases come both from constituents of a construction and the construction itself[3]

Particularly the ability to define arbitrary, partially specified constructions with a meaning contribution above and beyond their constituents will be relevant to the current discussion.

[3] Other common features less central here are a monostratal, non-derivational and non-modular architecture, see Fischer & Stefanowitsch (2006:3).

If we postulate an independent existence for complex structures, individual realizations of utterances can simultaneously instantiate multiple constructions along an axis of schematicity. For example, the following utterance contains multiple layers of form and meaning:

(3) *Give me the book!*

[V NP NP]$_{VP}$	(Ditransitive construction)
[*give* NP NP]$_{VP}$	(*give*-VP)
[*give me* NP]$_{VP}$	(*give-me*-VP)
VP$_{[+imperative]}$	(Imperative construction)

On the one hand, lexemes contribute meanings relating to giving, a book, and a recipient (indexical *me*). At the same time, it is not the form 'give' alone that signals imperative mood or ditransitive argument structure: *give* can be indicative, and can also govern a *to*-PP or other argument structures. Instead the configuration of the elements allows us to identify the multiple, concurrent constructions: imperative, ditransitive, etc. It is also likely that frequent combinations, such as *give me* are stored or 'entrenched' in the mental lexicon, and develop routinized idiosyncratic properties. These can include phonetic reduction (e.g. *gimme*, cf. Bybee 2013), but also pragmatic levels of meaning, such as the appropriateness and (im)politeness of combining *give*, the imperative and recipient *me* (for a third person we may have different pragmatic connotations, etc.).

The same holds for morphology, including compounding (cf. Booij 2009). If someone is a 'department head' or 'manager', we can say this in German with the following compound:

(4) *Abteilungsleiter* 'department head/manager'

[N-N]$_{N}$	(Unspecified N-N compound)
[N$_{[+ARG]}$-N]$_{N}$	(Argument structure compound)
[N-V*er*$_{N}$]$_{N}$	(Deverbal agent compound)
[N-*leiter*$_{N}$]$_{N}$	(*Leiter* compound, a type of 'manager')

While the interpretation of compounds like the above involves compositionality, it may well be the case that each of the constructions postulated above plays a part. The compound obeys the rules of N-N compounds at large (regular linking element *-s-* for modifier ending in *-ung*), but also shows argument saturation (the 'department' is the thing being managed). A one-to-one ability to convert verbal arguments to compounding argument structures does not obtain (see Gaeta & Zeldes 2012). Additionally the agentivity of the agent noun in *-er* is preserved, but this need not have been the convention for this construction. Other conceivable argument structure compounds are not licit in German, for example:

(5) **Die Dänemarkkönigin*
Denmark-Queen
'The Queen of Denmark'

The normal pattern in German for constructing monarch appellations, whose frame takes something like a RULED or GOVERNED argument, is using a PP with *von* 'of', e.g. *Die Königin*

von Dänemark 'the Queen of Denmark'. Note that there is no a priori way to know this, especially as an L2 learner (cf. Fillmore's 1977 'Innocent Speaker', who cannot know not to create the hypothetical compound). In some languages, monarchs are formed in exactly this way, e.g. using the Semitic Construct State construction in Arabic or Hebrew (see Borer 2008):

(6) Hebrew: $[N_{[+RULER]}\ N_{[+RULED]}]_N$

מלכת־דנמרק	[mal'ka-tˌdɛnmaʀk]	Queen-LINK-Denmark, lit. 'Denmark-Queen'
נשיא־גרמניה	[nɛ'si gɛʀˌmanja]	President-Germany, lit. 'Germany-President'

(7) Arabic: $[N_{[+RULER]}\ N_{[+RULED]}]_N$

ملكة الدنمارك	['malika-tˌl-danma:rik]	Queen-LINK-the-Denmark, lit. 'Denmark-Queen'
رئيس ألمانيا	[ra'ʔi:s ʔalma:nija]	President-Germany, lit. 'Germany-President'

Learners must learn which constructions are idiomatic and which are not, meaning that the acquisition of sub-patterns down to the level of 'compounds for monarchs' are likely to be part of the learning task. Saying the equivalent of 'the Queen of Denmark' with a prepositional phrase may be grammatical in Hebrew or Arabic, but it is not acceptable as the target norm.

If we accept the necessity of constructions at this level of specification for the acquisition of compounds, we quickly come to the line between theoretical and empirical questions about the grammar of compounds: which constructions can we identify? How do they differ in L1 and L2? To answer these questions, we need a concrete dataset, which will be presented next.

Method

Falko (*Fehlerannotiertes Lernerkorpus* 'Error-Annotated Learner Corpus') is the largest freely available corpus of L2 German writing (Reznicek et al. 2012). Data were collected at universities worldwide from advanced German learners and cover 39 different native languages,[4] allowing for broad typological comparisons. The proficiency level of the L2 writers was estimated as "advanced" using a C-Test (Klein-Braley & Raatz 1982) score above 60% (see Norris 2004: 269-309 for an in depth analysis and possible weaknesses of the test).

Falko contains multiple subcorpora, including essay data (Falko Essay), summaries (Falko Summary), a longitudinal corpus collected at Georgetown University (Falko GU) and supplementary essays collected in Britain (Falko WHIG; see Krummes & Ensslin 2014). For the Essay and Summary corpora, comparable L1 data exist (Falko Essay L1, Falko Summary L1), collected under the same circumstances and prompts. Since the summary corpora contain a high amount of technical terms taken over from the summarized texts and the longitudinal corpus contains heterogeneous data from various levels without comparable L1 data, I will focus on Falko Essay (V2.4) and WHIG (2.1). Prompts for essays followed those used by the International Corpus of Learner English (Granger et al. 2009), using questions such as: Are university degrees practical? Has feminism hurt women more than it has helped them? Or, does crime pay?

[4] I will use ISO 639-3 standard 3-letter language codes throughout; see Appendix.

The L2 data comprise some 250,000 word forms in 444 Essays, while the smaller L1 corpus covers over 68,000 word forms in 95 essays (see also Zeldes 2013 for numbers and results from earlier versions). All of the corpora were tagged automatically and lemmatized using the TreeTagger (Schmid 1994) and the STTS part of speech tag set for German (Schiller et al. 1999), and subsequently corrected manually with the aim of annotating the actual L2 text, including errors. This means that erroneous words are given tags corresponding to their apparent form, and not a normative one. At the same time, in order to capture information about errors, Falko contains a set of annotation layers called 'target hypotheses' (TH, or in German *Zielhypothesen*, ZH). The TH represent a version of what the annotator believes the learner is trying to say, but in correct German. The idea is that we cannot annotate for errors without having a notion of what we believe the correct form should be (e.g. an utterance like 'dogs runs' could be a verb or a noun error, depending on what we think the TH is). Although human annotators disagree considerably on the TH for specific sentences (see Lüdeling 2008 for an experiment), this form of annotation, coupled with carefully formulated guidelines and annotator training, has led to considerable success in investigating properties of learner language (see Reznicek et al. 2013).

Figure 1 illustrates annotations for an L2 fragment "*als das, was im Buch steht zu unterrichten*", which means '(rather) than to teach what it says in the book'. The learner has chosen a word order that is not completely ruled out in German, though orthographic rules require a comma to appear before *zu unterrichten* 'to teach' in this case. A native annotator preferred the more natural sounding order "*als das zu unterrichten, was im Buch steht*", literally 'than that to teach, what in the book says'.

ZH1	als	das			,	was	im	Buch	steht	,	zu	unterrichten	.
ZH1Diff										INS			
ZH1lemma	als	d			,	was	im	Buch	stehen	,	zu	unterrichten	.
ZH1lemmaDiff										INS			
ZH1pos	KOKOM	PDS			$,	PRELS	APPRART	NN	VVFIN	$,	PTKZU	VVINF	$.
ZH1posDiff										INS			
ZH2	als	das	zu	unterrichten	,	was	im	Buch	steht				.
ZH2Diff			MOVT	MOVT							MOVS	MOVS	
ZH2S	s5												
ZH2lemma	als	d	zu	unterrichten	,	was	im	Buch	stehen				.
ZH2lemmaDiff			MOVT	MOVT							MOVS	MOVS	
ZH2pos	KOKOM	PDS	PTKZU	VVINF	$,	PRELS	APPRART	NN	VVFIN				$.
ZH2posDiff			MOVT	MOVT							MOVS	MOVS	
tok	als	das			,	was	im	Buch	steht		zu	unterrichten	.

Figure 1. Target hypothesis (TH) layers for an L2 utterance fragment.

The layer labeled ZH1 (=TH1, first target hypothesis) offer a minimal correction of the text which the annotator feels is grammatically acceptable (see Reznicek et al. 2012 for detailed guidelines). Here only the comma has been introduced, and its insertion is marked in ZH1Diff with the tag 'INS'. Other tags include 'CHA' for changing the form or 'DEL' for removing an item, and separate levels indicate edits to the part-of-speech or lemma, rather than the word form (ZH1pos/ZH1lemma). ZH2 contains more radical changes. This layer is meant to give what

annotators feel is a natural sounding equivalent expressing the meaning intended by the learner. Here *zu unterrichten* 'to teach' has been moved, as signaled by the tags MOVS (move-source) and MOVT (move-target). By its nature, this layer is more prone to variant annotations by different annotators. Beyond TH1 and TH2, there are further layers, including a special layer for verb errors (ZHVerb) and a TH0, which is identical to TH1, but without changes to word order. In this study I will only use TH1 and TH2.

A further form of annotation used in the Essay corpus (but not yet in WHIG) is automatically generated dependency trees, using the parser described in Bohnet (2010). Although these are not manually corrected, they can be useful in recovering argument structure relations, as we will see in Section 4. Figure 2 illustrates such trees, which give the grammatical relation of each word to the word that governs it. Dependencies were parsed on the TH1 layer to increase accuracy, but can be projected back to the original text via scripts using the ZH1Diff annotations.

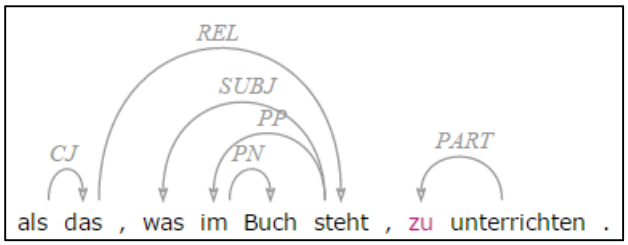

Figure 2. Dependency parse for the fragment from Figure 1.

All of the annotation layers described here can be accessed freely using the ANNIS interface (Krause & Zeldes 2014) at http://korpling.german.hu-berlin.de/falko-suche/ and the corpus can be downloaded for offline use from the corpus website at: https://www.linguistik.hu-berlin.de/en/institut-en/professuren-en/korpuslinguistik/research/falko/.

A caveat for using Falko for this study is that compounds are not annotated in the corpus directly. In order to extract information about compounds, all items tagged as nouns (NN) were compared to large lists of simplex nouns derived from the lexicon of the SMOR computational morphology for German (Schmid et al. 2004) and a large corpus with compound lemma tagging (TuePP-D/Z, Müller 2004). For nouns not found in the lists obtained in this way, an automatic compound analyzer was written which attempted to break up words into compounds comprising the longest possible head from the list (on the right-hand side), possible linking elements, and modifier(s) to the left which had to be phonologically possible German words (e.g. not non-syllables such as *Zw-*) and not on a list of known prefixes (e.g. *Be-*, *Un-*). Approx. 10% of candidates could not be identified with high certainty, and these were corrected manually on the right, and cases that were judged not to be nouns at all (tagging errors) were removed. Table 1 gives an overview of the results of this process.

Table 1. Distribution of compounds and simplex nouns in the L1 (N=95) and L2 (N=444) essays.

	automatic				corrected		
	L1	L2	Total		L1	L2	Total
simplex	8,240	30,279	38,519	*simplex*	9,077	33,306	42,383
compound	1,453	3,546	4,999	*compound*	1,658	4,355	6,013
unknown	1,130	4,216	5,346	--			
Total	10,823	38,041	48,864	Total	10,735	37,661	48,396

Results

Errors in L2 Compounds

In approaching the analysis of errors in the L2 data, we may consider several questions: is compounding difficult or error-prone for the learners? What proportion of errors will we find and of what kind? Will these errors be directly linked to the compound formation process, or will they be the same errors we find throughout L2 data? This last question is tightly linked to our ability to understand how compounding works in L2 grammar and demands a classification into compounding-related and non-compounding-related errors, which I will distinguish below.

Using TH1 data, which contain all noun forms ruled out as grammatical by annotators (but not corrections for natural style, logic/rhetoric), we can find forms of compounds that would have been adjusted by a native speaker, and consequently contain errors. In the 4,355 L2 compounds identified in Table 1, 483 errors can be found using manually checked queries for TH1 changes, a ratio of 11.1% (cf. similar results in Zeldes 2013, p. 249). Table 2 gives the proportions of major error types, classified manually (since one compound may contain multiple errors, the total exceeds 483).

Table 2. Overview of error type in the L2 compounds, including compounds corrected by annotators via change to a simplex noun.

	Frequency	% Compounds	% Errors
Derivation	10	0.002	0.020
Linking	172	0.039	0.339
Inflection	139	0.032	0.274
To simplex	11	0.003	0.022
Lexical	38	0.009	0.075
Modifier	68	0.016	0.134
Head	70	0.016	0.138
Total	508	N/A	100

As Table 2 shows, the majority of errors are not related to compounding per se: spelling errors in heads and modifiers which can and do appear for the same lexical items outside of compounds, inflectional errors, which generally affect the head in the same way as a simplex

noun, and derivation errors creating an unacceptable head/modifier. The following are examples of these error types, which will not be discussed further, along with similar simplex errors:[5]

(8) Derivation: *Arbeits|such<u>ung</u>* 'job search' (the *-ung* suffix is an erroneous derivation from *suchen* 'to search', TH: *Arbeits|such<u>e</u>*; L1=pol), cf. simplex error *Train<u>ung</u>* (for *Training*, L1=eng)

(9) Inflection: *Dieb|st<u>a</u>hl<u>e</u>* 'thefts', lit. 'thief-stealings' (TH: *Dieb|st<u>ä</u>hl<u>e</u>*; L1=eng), cf. similar simplex plural errors: *Hauser* 'houses' (TH: *Häuser*; L1=nde)

(10) Modifier and head spelling: *Bach<u>er</u>l<u>e</u>r-Fak<u>a</u>ltät* 'Bachelor Faculty' (TH: *Bachelor-Fakultät*; L1=jpn), cf. Simplex *Bache<u>ll</u>or* (L1=ell)

By contrast, compounding-related errors are of central interest here, since they cannot affect non-compounds and are thus intrinsically linked to the acquisition of compounding as a construction or family of constructions. For German, these are errors relating to compound morphology, i.e. the use of correct linking elements; and compounding syntax/semantics, regarding the correct combination of head and modifier in a native like way. Another special type of possible error is using a compound where no compound is appropriate, i.e. compounds that a native speaker would prefer to correct to simple nouns.

The most frequent category both within compounding-specific errors and in total is the linking element category, amounting to 172 occurrences, or about 34% of errors in compounds, followed by the much less frequent, though very interesting errors related to lexical blocking and combination errors. Linking errors can be subtractive (element missing), additive (element superfluous), or both (i.e. substitution of a correct element by an incorrect one):

(11) Subtractive: *Arbeit|markt* 'job market' (TH: *Arbeit<u>s</u>|markt*; L1=eng)

(12) Additive: *Abend<u>s</u>|essen* 'dinner', lit. 'evening food' (TH: *Abend|essen*; L1=eng)

(13) Substitution (both): *Geschlecht<u>en</u>|rollen* 'gender roles' (TH: *Geschlecht<u>er</u>rollen*; L1=eng)

Although this is a large proportion of the errors, it should be noted that in total less than 4% of compounds exhibit linking element errors, which shows that (advanced) learners have good command of the rules involved (see Hartkamp & Schneider-Wiejowski 2010 on the variability and occasional unpredictability of linking rules). The distribution of linking element errors is given in Figure 3, which shows very strong trends, partly related to the frequencies of the elements themselves.

[5] I will use vertical lines to separate compounds in German for clarity where necessary; these were not present in the learner data. Errors and their corrections have been underlined.

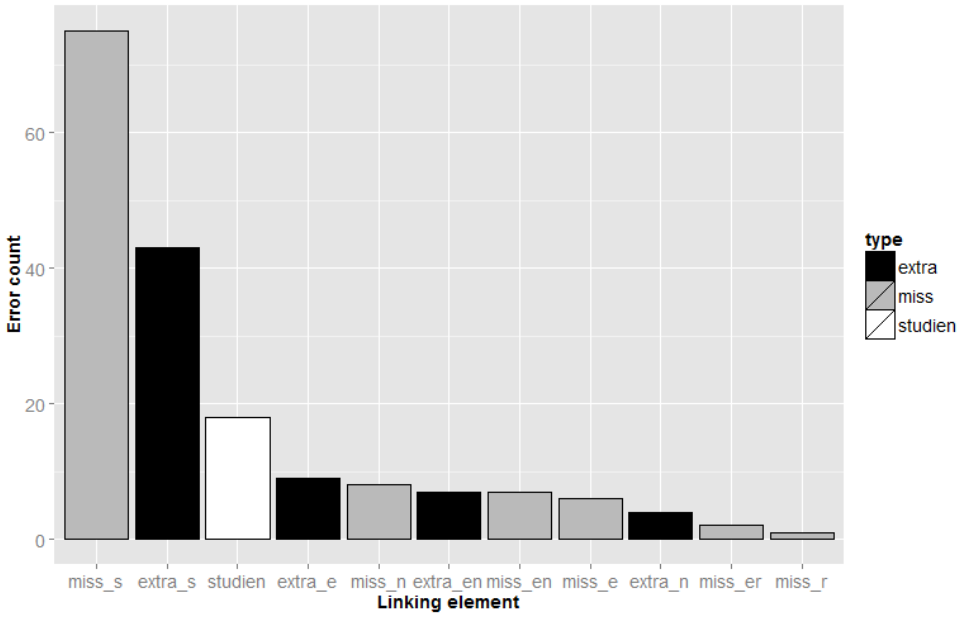

Figure 3. Linking element error types, for missing (grey) and superfluous (black) linkers. The exceptional *Studium-* for *Studien-* is in white.

As the figure shows, errors involving the linking *-s* are by far the most common. This is perhaps not surprising, since it is not only common but also the most variable linker in native usage, showing inconsistent realization in many compounds in usage (see Michel 2009, Nübling & Szczepaniak 2011).[6] Possible covariates to *-s* variation known from L1 literature could not be identified in the L2 data, e.g. unstressed prefixes or non-trochaic/monosyllabic structure on the modifier (Nübling & Szczepaniak 2011) or morphological complexity/presence of derivational affixes (Kürschner 2005). The only pattern apparent in the L2 data is the preponderance of modifiers ending in *-t* for the missing *-s* cases. Modifiers in *-t* form 62% of the linking error tokens, but of these almost two thirds of cases with *-t* are missing the element. The confusion involving *-t* may have perceptual reasons (cf. Goldschneider and DeKeyser 2001; salience in the transition from [ts] may be low, and in some cases the head also begins with *s-*, meaning the error is purely orthographic), but partly it can be ascribed to conflicting evidence in such common types as *Arbeit<u>s</u>|erfahrung* 'work experience' versus *Projekt|arbeit* 'project work'. For nouns such as *Arbeit* 'work' we also find forms that normatively exhibit no linker, e.g. *Arbeitgeber* 'employer, lit. work-giver'. Learners may have a difficult time abstracting a rule from such conflicting cases due to low cue reliability (cf. MacWhinney 2012; I thank Lourdes Ortega for commenting on this issue), and for some lexemes memorization is the only solution.

Although missing elements are more common than superfluous ones overall, the latter are not uncommon (a ratio of 99:63), meaning learners are generally aware of the necessity of linkers,

[6] Although *-s* is the most frequent overt linker (estimated at 17-25% of cases, see Nübling & Szczepaniak 2013), this alone cannot account for Figure 3 – the runner up, *-en*, has 11-15% frequency in previous studies and shows very few errors, and zero-linking is the most common option overall (around 60%, ibid.).

even if they have not internalized the correct form or rule. Additionally, the neoclassical modifier for *Studium-* 'study/studies' forms a common and special case, since it should correctly be realized as *Studien-* in compounds. Its frequency is a direct result of one of the essay topics, about whether or not university degrees are practical enough (optionally it could be added to the 'missing -en' category, which would boost its frequency, but the two have been distinguished here). It is therefore clear that the general principle of linking has been acquired in the vast majority of cases, but specific instructional attention, especially for *-s* and with modifiers in *t-*, could help to eliminate over 65% of the errors.

Much less frequent than linking errors is the category of lexical errors, often related to lexical blocking, which are discernable through changes on the TH1 layer. These types of errors are particularly interesting, since they attest to exceptions to the general acceptability of abstract rules of the type N+N=N which learners are meant to acquire for productive use. Below are examples that annotators felt the need to correct:

(14) *oft befindet sich das Studiumort nicht in der Heimstadt*
'often the place of study is not in the hometown' (TH1=*Heimatstadt*; L1=ukr)

(15) *ein Kampf nach Menschlicherechte* 'a fight for human-rights' (TH1=*Menschenrechte* 'lit. person-rights') (L1=fra)

These examples are likely due to loan translations such as *hometown = home+town* (*Heim+Stadt*), which is outwardly well formed but lexically blocked (cf. Rainer 1988) by the lexicalized *Heimatstadt*. The example of 'human rights' has 'human' as an adjectival modifier in **Menschlicherechte*, but should be an N+N compound with the noun for human, *Mensch* 'person, human'). Other cases show use of a less appropriate lexeme despite no clear lexicalization, or use of infelicitous loans:

(16) *Tenniswoman* 'female tennis player' (English loan – TH1: *Tennisspielerin*; L1=fra)

(17) *Universitäatskursus* 'university course' (Latin loan – TH1: *Universitätskurs*; L1=fra)

Loan translations can also take on a Germanized morphology, but can still be superseded by lexicalized alternatives:

(18) *Nativsprecher* 'native speaker' (TH1: *Muttersprachler*; L1=ces)

(19) *Chancespielen* 'game of chance' (TH1: *Glücksspiele*; L1=fra)

These types of errors were very rare, affecting less than 1% of compounds, though they only include those cases that were ruled out as strictly unacceptable. Using TH2 we can find further cases that are morphologically well formed, but odd, such as:

(20) *Der Beruf, den er ausübt lässt mehr oder weniger Spuren auf dem Lebensblatt der Gesellschaft.* (L1=ron)
'The profession that he practices leaves more or less traces on the life-page of society'

Cases like these are usually loan translations, in this case from Romanian *pagina vieţii* 'page of life', which are idiomatic in the L1 but not in the target language. They nevertheless exhibit learners' ability to map two German lexemes not heard together before onto the productive $[N\text{-}N]_N$ compounding construction, even though the transferred construction does not correspond to the target one (in Romanian *vieţii* is a free genitive). Using TH2 we can also find cases of compounds better realized as syntagms and vice versa, i.e. 'missing' compounds (which may partially account for the lower compounding rate in L2 data, see below) and 'superfluous' ones:

(21) *Oft in Zeit von Kriegen* 'often in time of wars' (TH2: *Kriegszeiten* 'war times'; L1=eng)
(22) *Ob es mit Gefängniszeit oder eine schlechte Gewissen endet* 'whether it ends with jail time of a guilty conscience' (TH2: *Gefangenschaft* 'imprisonment'; L1=eng)

These cases are qualitatively very interesting, but overall very rare and therefore difficult to group into meaningful classes.

Native Compounds versus Learner Compounds

Beyond the correctness of the forms they produce, we would also like to know whether learners acquire the characteristic 'compounding-happiness' of the L1 target variety. Table 1 already reveals that learners form significantly fewer compound tokens than natives (L1: 15.4%, L2: 11.5%, $p<2.2e\text{-}16$, $\chi2=115.29$, N=48,396). However for this difference to be interesting as a property of L2 German as its own interlanguage, we must perform a contrastive analysis of learners with natives as groups of individuals (Contrastive Interlanguage Analysis, cf. Granger 2002). As the boxplots in Figure 4 reveal, differences in compounding frequency are pronounced (L2 median just under L1 interquartile range), but there is a substantial range of outliers for L2.

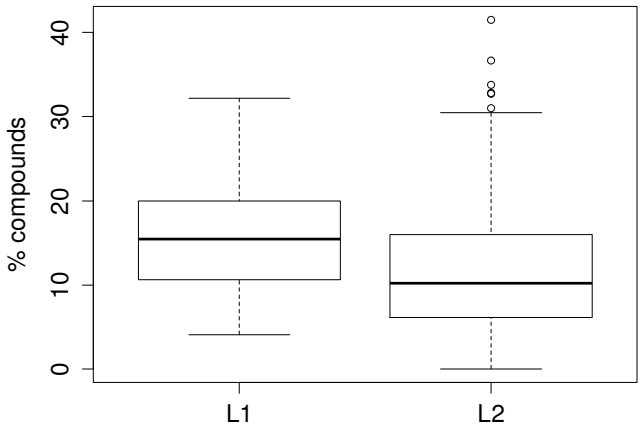

Figure 4. Boxplots for compound proportion per writer in the essay corpora.

There is some expectation that learners may prefer analytic constructions to more morphological constructions (cf. Klein & Perdue 1997; I thank an anonymous reviewer for commenting on this). Are learners in fact using alternatives to compounding to express nominal modification, which would explain their lower frequency? Are there multiple subgroups among the learners? If so, what distinguishes them? Is there an L1 transfer effect? Or differences due to proficiency?

In terms of alternative modifier constructions, it seems that learners underuse virtually all of the main alternative competitors to compounding, including adjectival, clausal and prepositional variants. Table 3 shows the frequencies of adjective modifiers (ADJA), genitive modifiers (AG dependency), relative clauses (PRELS) and prepositions (APPR[ART]), further broken down into *von* 'of' and others (WHIG data have been excluded since parses are not available).

Table 3. L1 and L2 frequencies of constructions competing with compounds.

	f(L1)	*norm(L1)*	*f(L2)*	*norm(L2)*	*ratio*	*significance*
ADJA	2753	0.0402	5052	0.0411	+2.34%	ns
GMOD	942	0.0137	1569	0.0127	-7.11%	ns
PRELS	666	0.0097	991	0.0080	-17.02%	***
APPR(ART)	4766	0.0695	8092	0.0659	-5.31%	***
…von	410	0.0059	797	0.0064	+8.41%	ns
…¬von	4356	0.0636	7295	0.0594	-6.60%	***

* p < .05; ** p < .01; ***p < .001

Learners' use of adjectives is not significantly higher to compensate for the missing compounds, and underuse of all other possibilities is apparent, with the sole exception of the preposition *von* 'of', suggesting that 'missing' compound modifiers are not compensated for, but simply omitted. This results in less modification overall in the L2 data (cf. Hirschmann et al. 2013, Vyatkina et al. 2015 for similar findings on adverbial, adjectival and prepositional modification).

To answer the questions about proficiency groups and L1s of learners, we must take care to get reliability estimates for the differences between subgroups and the significance of correlations. Figure 5 lays out the average ratio of compound nouns per text in each L1 group for the largest L1s in the corpus (the remaining languages had fewer than 5 documents each). A simple expectation would be for languages with less compounding to exhibit fewer compounds: this is borne out for the Romance languages in white and Greek next to Italian, all on the right side of the plot.

For Slavic and Germanic, the picture is more complex: although Norwegian and Danish speakers actually produce slightly more compounds on average than natives (difference still within the margin of error), Afrikaans is significantly lower down, and Dutch is at the end, while English takes a middle position (all significantly below L1 German controls, the column "deu"). Similarly Slavic and Turkic are not contiguous, with Ukrainian nearly the same as L1 controls but Russian significantly below. Chinese ("zho"), which has productive and recursive compounding, does not differ significantly from German due in part to a smaller dataset and a larger margin for error (only 7 texts, meaning the average figure should be taken with a grain of

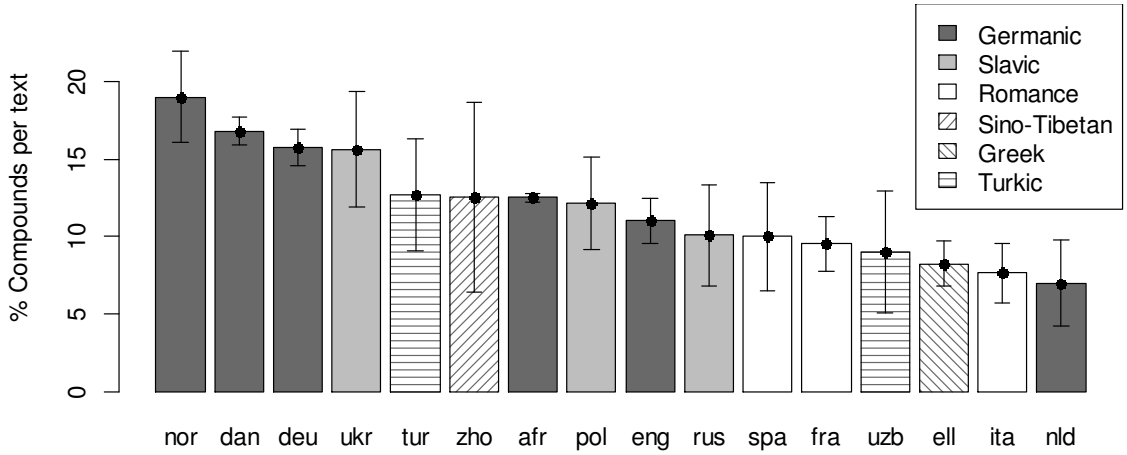

Figure 5. Average compounding ratios per text with standard error for the top 15 L1s and German (deu).

salt). The picture that emerges is that writers from L1s low in compounding (see Lieber & Štekauer 2008 and Gaeta & Schlücker 2012 for discussion) are also low in L2 German compounds, but those with L1s rich in compounding do not necessarily exhibit prolific use of compounds.

Beyond L1 interference, another suspect factor potentially influencing compound frequency is proficiency. Is the ability or tendency to use compounds related not only to L1 background but also to proficiency in this sample of very advanced learners? And if so, is productive L2 compounding really native-like in more proficient learner data? To evaluate the proficiency-based hypothesis, we can make use of C-Test scores as a rough estimate of learner competence in written German. The advantage of relying on C-Test scores to gauge proficiency is that they are entirely independent of the text. Text-based metrics such as mean sentence length and text length, syntactic complexity and others are sometimes used as indices of L2 proficiency (see Norris & Ortega 2009 and Lu 2010 for an overview and discussion), but any such objective metrics might be intrinsically related to compounding itself (e.g. more compounds may mean fewer words per sentence, etc.). As Figure 6 shows, C-Test averages (N=444, Ø=80.24, sd=9.97, range 60,100) do not offer an explanation for the L1-based differences found in Figure 5.

Although L1s with a high mean rate also have high C-Test averages (nor, dan, ukr), other L1s with high C-Tests exhibit lower compounding rates. Looking at individual learners on the right panel, we see a better, but weak correlation (rho=0.24, $p<0.001$) between compounding and C-Tests. The upward spread of the data to the top-right suggests the same as the L1-based evidence: A high compounding rate is indicative of high C-Test scores to some extent, but *not* using many compounds does not tell us as much. Since compounding is a combinatory process, it makes sense that, like other measures of complexity, it will correlate with proficiency. In particular, in many scenarios adding a compound modifier is optional for a given communicative intent (we may speak of learning 'languages' *Sprachen* or specify 'foreign languages', compound *Fremdsprachen*), and learners are known to forgo optional modifiers in general (Hirschmann et al. 2013).

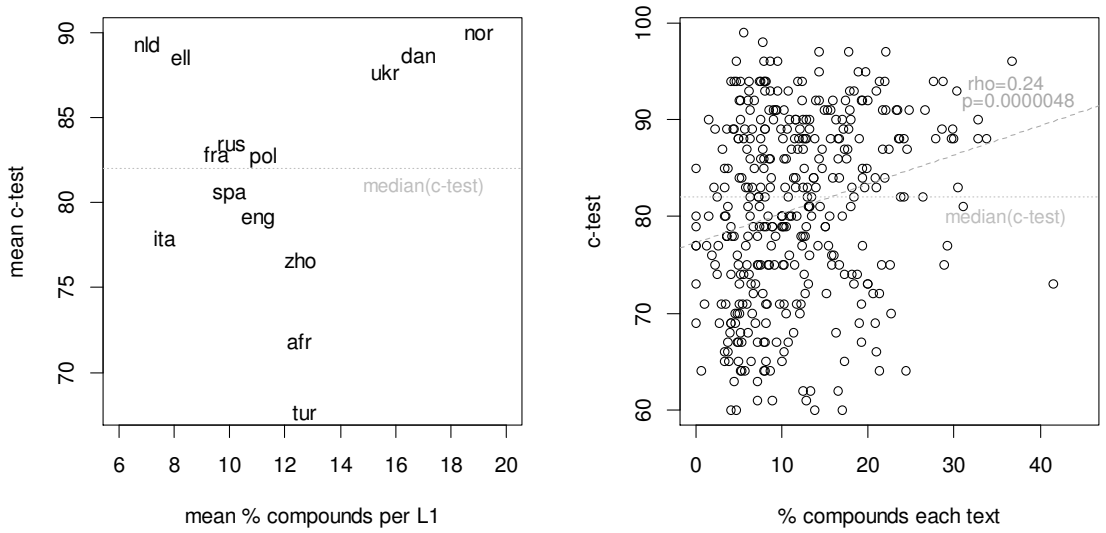

Figure 6. No correlation between compounding per L1 and C-Test; weak correlation between C-Test and individual learner compounding.

This interpretation of lower compounding frequency suggests a connection with lower productivity – the ability to use the compounding system dynamically to subtly modify existing meanings. The next question in interpreting differences between L1 and L2 compounds, and particularly differences that may be informative of the status of compounds in learner grammars, should therefore be: are learners using compounds productively in a similar way to natives, or are the bulk of compounds in the data above lexicalized items, which tell us less about compounding as a process? The remainder of this chapter is dedicated to this question.

Compounding and Productivity

Measuring productivity is difficult, since we can never be sure whether a certain form is being generated by a learner productively as a neologism, or simply being repeated from memory (see Bauer 2001:34). By analogy to studies of productivity in morphological derivation (e.g. showing that affixes like *-ness* are more productive than *-ity* or *-dom*), we can estimate productivity for compounding by looking at total vocabulary exhibited (notated V: more vocabulary types means more productive) and the proportion of rare items, which are taken to be a superset of true neologisms (see Baayen & Lieber 1991, Baayen 2001, 2009). In particular, so called hapax legomena, items with a vocabulary frequency of 1 (or 'V1') and their proportion in a sample have been taken as indicative of the rate of saturation or repetitiveness with which a process is used. Although these measurements were originally developed to study the variety of stems found with different affixes, they can also be applied to the combination of free stems and syntactic processes (Zeldes 2012), which makes them relevant for the study of compounding productivity.

Taking the L1 and L2 populations each as a whole, we can expect two conflicting tendencies: 1. Learners are generally taken to be less productive than natives, more repetitive, and their

vocabulary is by nature initially smaller (cf. Ellis 2002, 2012; Milton 2009).[7] 2. Learners are not constrained as strongly by the rules of the target language and may produce a wider variety of types due to errors, lack of lexical blocking, etc. Figure 7 charts vocabulary growth curves (using the library zipfR in R, Evert & Baroni 2007) for compounds in the entire L1 and L2 datasets, with growing sample size N on the x-axis and the number of types encountered at that point V on the y-axis. Since it becomes progressively harder to find novel items as the sample grows, a fair comparison can only be drawn at the largest possible common sample size (see Gaeta & Ricca 2006), dictated by the smaller L1 dataset (1,658 compound tokens in 95 essays).

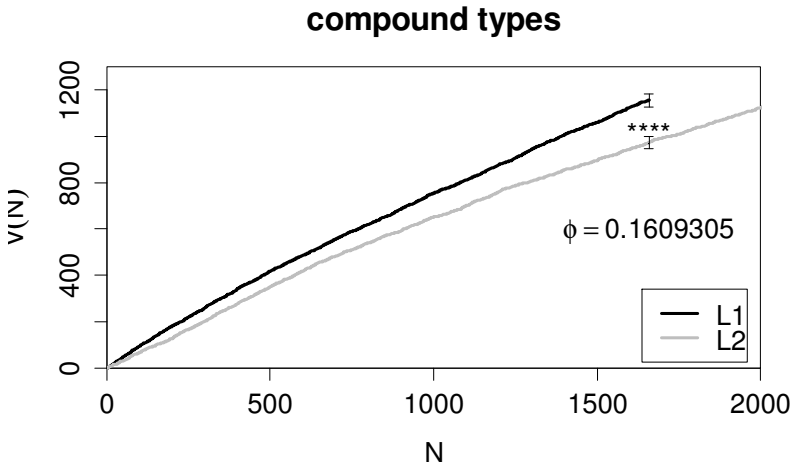

Figure 7. Vocabulary growth curves for L1 and L2 compound types.

As the figure shows, L2 data has a smaller vocabulary, meaning that even with the presence of erroneous compounds, the lower productivity of L2 data leads to fewer types in total. This is despite the fact that some words have multiple spelling types included, e.g. the modifier *Universität* 'university' appears in L2 data as *Universitäat*, *Universitäs*, *Universisität* and more. The difference in V at the common sample size is of 182 types, so that L2 has almost 16% less vocabulary than L1 in total, a very significant difference (p= 5.645e-11, ϕ=0.1609).

Although this tells us something about the learners as a whole, it does not tell us how individual learners vary in terms of productivity. For example, is proficiency as estimated by the C-Test related to productivity? Taking vocabulary measurements from single texts is not possible, since a short text does not allow for so much repetition: almost all compounds will have single mentions, whether or not they are lexicalized, and the dependency of V on sample size (Gaeta & Ricca 2006) will mean that we are limited to the shortest possible text. A possible solution to this problem is to place multiple learners, for example ones with similar C-Tests, into the same 'bin' and collect vocabulary data on each group.

[7] Vocabulary refers here to production vocabulary, and not the available receptive vocabulary, which cannot be assessed using the corpus (cf. Nation 1990, de la Fuente 2002).

To examine the relationship between proficiency and productivity, Figure 8 gives V and V1 measurements for equal sample sizes taken from bins grouping together each quartile of the C-Test scores (bottom 25%, lower middle, top middle and top 25% of scores).

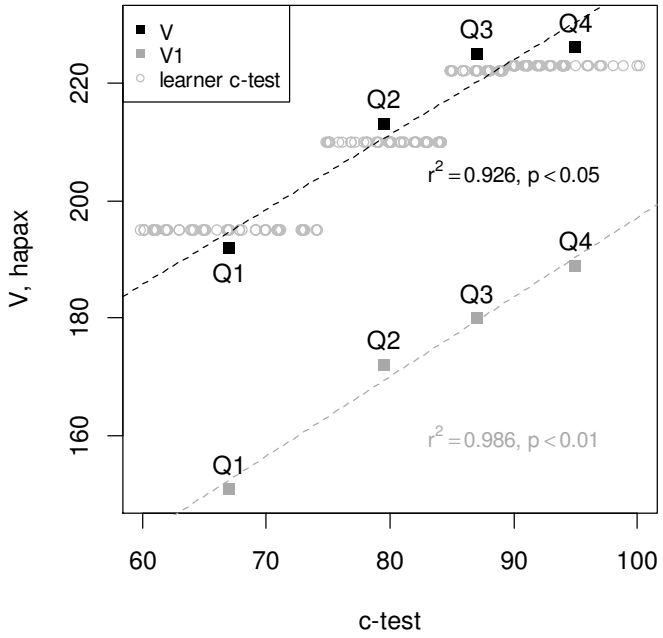

Figure 8. V and V1 in binned C-Test quartiles.

As the data show, there is a clear correlation between vocabulary size in each bin and the C-Test quartile: there is a steady and significant progression from Q1 to Q3 in V, with Q3 almost identical to Q4 (pairwise difference insignificant). However in V1 (unique types), we see the rise through all four bins, with an even stronger and more significant correlation. In other words, from Q3 to Q4 in C-Test scores there is no longer a rise in total vocabulary, but nevertheless a rise in the proportion of items that are rare within that vocabulary. This corresponds to known facts about overly skewed distributions in L2 data (the 'teddy bear effect', Ellis 2012), which seem to be mitigated for the highest C-Test performers. Whereas lower proficiency learners produce vocabulary type distributions dominated by a handful of favorite types (the 'teddy bears') and relatively few rare items, more proficient learner show a more nativelike spread, in which the greater majority of types is rare (many unique items), and the relative dominance of the prototypical items is less overpowering.

A final point for examination is whether the individual types of novel compounds that learners produce are similar to native compounds. As I argued earlier, target language usage of compounding requires access to lexicalized and partially lexicalized compounding constructions. Some of these take the form of more or less abstract compounding semantics or morphosyntactic constraints, while others involve concrete heads and modifiers (extensible 'positional families' of compounds, which also have consequences for faster processing of novel forms, see Krott et al. 2007). Consider for example the following abstract constructions giving conventionalized compounding semantics in German, and a partially filled modifier pattern (the notation follows

Booij 2010 and gives a constructional schema on the left with its corresponding stored meaning on the right):

(23) *Purpose compound:* $[N\text{-}N]_N$ ↔ |*N for N*|, e.g. *Arbeitstisch* 'work table'
(24) *Agent compound:* $[N\text{-}[V]er_N]_N$ ↔ |*one, who Vs N*|, e.g. *Büroleiter* 'head of the office', 'one who heads/leads the office'
(25) *Modifier compound:* $[Profi\text{-}N]_N$ ↔ |*professional N*|, e.g. *Profi-Programmierer* 'pro programmer'

Much like the convention for not expressing heads of states by compounds in German (but in Semitic languages), these are normal and routinized ways of forming German compounds, or what may be thought of as 'mini-constructions', following Boas (2003). Many novel examples are modelled after prototypes such as those in (23-25) and pick out the same aspects of the semantics of each constituent. For example, the word *Profi-Programmierer* 'pro programmer' designates a programmer who is professional, blocking an alternative interpretation using the agent compound construction. To see the alternative reading, consider the following example:

(26) *Experten-Programmierer* ↔ *|*programmer who is an expert*|
#|one who programs experts|

Although the second reading is odd, the outward form $[N\text{-}[V]er_N]_N$ and the lack of an entrenched construction $[Extperten\text{-}N]_N$ enable the agent compound reading. The same reading is highly odd for ??*Profi-Programmierer* 'one who programs pros'. Note that the *Profi-* 'pro' construction is non-trivial to induce: a corresponding compound e.g. in Hebrew or Arabic is not licit (Heb. compound **metaxnet-ha-miktso'an* 'the pro-programmer'; an N+Adj construction with article agreement is required instead: *ha-metaxnet ha-miktso'i*). Learners successfully employ such modifier constructions (e.g. *Profi-akademiker* [sic] 'pro-academic', *Profi-Basketballer* 'pro basketball player', *Profi-Sportler* 'pro athlete', etc.) and alternatives of the ?*Experten-* type are not attested.[8]

With the notion of subtypes of compounding constructions in mind, we can now ask whether the learners use the same constructions as natives and whether or not there are differences in frequency. In order to focus on the most productive examples, rather than lexemes that can be learned wholesale, Table 4 gives 200 randomly selected hapax legomena from each group, classified by compound paraphrase (translations are approximations).[9] For each class, results are divided between argument structure compounds (arg) and non-argumental modifiers (mod).

Table 4. Samples of 200 hapax compounds from the L1 and L2 data divided by subtype.

[8] There are many other examples of head and modifier mini-constructions, e.g. *Gesamt-* 'total-', *Haupt-* 'main', *Riesen-* 'mega-, huge-', but these cannot be discussed for space reasons; see Zeldes (2013).
[9] Many suggestions exist on how to subdivide compounding semantics (Levi 1978, Fanselow 1981, Motsch 2004:376-420), and famously no exhaustive list is possible (see Downing's 1977 'apple juice chair'). The use of PP paraphrases is meant to offer an intuitive rough classification without imposing a strong theory on observations, see Gaeta & Zeldes (2012).

L1				L2			
type	mod	arg	total	type	mod	arg	total
adjectival	28	1	29	*adjectival*	31	0	31
accusative	0	28*	28*	*accusative*	0	13*	13*
PP an 'toward'	2*	1	3*	*PP an*	12*	1	13*
PP auf 'upon'	1	3	4	*PP auf*	2	5	7
PP aus 'out of'	5	0	5	*PP aus*	2	0	2
PP bei 'at'	1	1	2	*PP bei*	2	0	2
PP für 'for'	33	0	33	*PP für*	25	3	28
PP in 'in'	13	3	16	*PP in*	12	0	12
PP mit 'with'	2	1	3	*PP mit*	4	0	4
PP nach 'after'	1	0	1	*PP nach*	0	2	2
PP über 'over/about'	0	1	1	*PP über*	0	2	2
PP von 'of'	34**	28*	62	*PP von*	58**	14*	72
PP wegen 'because'	1	0	1	*PP wegen*	4	0	4
PP zu 'to'	1	2	3	*PP zu*	0	1	1
copular	0	0	0	*copular*	2	0	2
(other)	7	2	9	*(other)*	5	0	5
total	129	71***	200	total	159	41***	200

* p < .05; ** p < .01; ***p < .001

Some significant differences are apparent: Learners have a high proportion of compounds that can be paraphrased by *von* 'of' (e.g. *Arbeitwelt* 'work world, world of work'), but only half as many *von* compounds expressing argument structure (native deverbal *Verdienstausfall* 'loss of income'). Similarly, the proportion of accusative synthetic compounds is significantly lower (native *Kartenverkäufer* 'ticket seller', L2 *Strassensauberer* [sic] 'street cleaner'), and some of these go back to fully or partially lexicalized head patterns, e.g. *Abteilungsleiterin* 'department manager (fem.)', with a feminine derivation of the very common head pattern [N-*leiter*]$_N$ 'N manager'. Learners exhibit more modifier and less argument compounds in almost every category, and in all significant cases. The most balanced, target language near categories, are locational and purpose compounds with *in* and *for* paraphrases, e.g. *Betreibspraktikum* [sic] 'practicum in a business' or *Bewerbungsbrief* 'application letter, letter for an application'. It seems these cases, where compound semantics are more ad hoc than a morphosyntactic requirement of the head, are more accessible to the learners.

Conclusion

The results found in this study have demonstrated several important facts about how advanced learners of German use compounds in written language: Learners are less productive, more repetitive than natives and exhibit some errors. At the same time, error rates are overall relatively low, confirming previous findings (Heidermann 1997:58-60, Fandrych & Tallowitz 2008:202) to the effect that compounds are acquired early and well as a means of expanding vocabulary. Learners seem to have access both to highly abstract patterns at the level of [N-N]$_N$ and [N-V*er*$_N$]$_N$ but also to partially filled positional families of the type [N-*leiter*]$_N$ 'N manager' or [*profi*-

$N]_N$ ‘pro N’, and of course also to lexicalized compounds. Deviations from licit prototype-based constructions are extremely rare and usually due to transfer errors.

Although compounding is generally rarer in L2 data, some of the highest compounding rates are apparent in L1s with productive compounding, such as Norwegian and Danish. Some of the least ‘compounding happy’ L1s (especially Romance) exhibit very low compounding rates, suggesting a transfer phenomenon is involved. At the same time, not all productively compounding languages produce learner texts with a substantial proportion of compounding (notably Dutch and Chinese), and a better correlation can be achieved between learner proficiency as evaluated by a C-Test and the prevalence of compounds. This is especially true in one direction: learners who exhibited many compounds tended to have high C-Test scores.

As the analysis of rare compounds and qualitative examples in the data reveal, what distinguishes advanced learners from natives, and what improves along the axis of proficiency, is less the inventory of compounding constructions, and more the degree to which they are used productively. While vocabulary grows steadily into the upper half of C-Test scores, a productive skewed distribution with a high proportion of rare types is progressively more apparent with a differential even between the upper quartiles. A special barrier seems to affect productive use of argument structure compounds in our sample: learners are more comfortable adding a relational modifier of the *von* ‘of’ kind and less inclined to realize verbal valency in compounds. Arguably, compounds require higher proficiency, because the connection of head and modifier is less specified and relies on proficiency with the corresponding verb frame, which may not be a given.

The findings of this study suggest that there may be substantial didactic benefits to devoting attention to compounding patterns that are underused in the data explored here. With the learning goal of ‘advancedness’ (Byrnes 2012) in mind, the usage data indicate that argument structure compounds should be addressed more in the advanced German curriculum, with particular attention to the common type with agent noun heads in deverbal *-er*. The relationship between verbal and nominal(ized) argument structure appears to be a barrier, which needs to be crossed to produce target language like nominalizations that do not rely on paraphrases with *von* ‘of’. In the area of morphological errors, it is clear that there is a parallel between the most unpredictable linking element alternation between *s-* and zero in L1 and the most frequent error type in L2. Issues of saliency at the compounding boundary are also indicated by the data. Finally, the formation of productive compounds, above and beyond the acquisition of common established types, is part of what characterizes high level advancedness in L2 German, and therefore merits attention in teaching as well. For this purpose, genuine examples from corpus data can serve the added role of giving authentic interesting examples.

Acknowledgments

I would like to thank Lourdes Ortega, Ande Tyler and an anonymous reviewer for valuable comments on previous versions of this chapter, as well as Marc Reznicek und Heike Zinsmeister for their input on earlier stages of this work. The usual disclaimers apply.

Appendix: ISO 639-3 three letter language codes

Afr=Afrikaans	eng=English	nor=Norwegian	swe=Swedish
ces=Czech	fra=French	pol=Polish	tur=Turkish
dan=Danish	ita=Italian	ron=Romanian	ukr=Ukrainian
deu=German	jpn=Japanese	rus=Russian	uzb=Uzbek
ell=Greek	nld=Dutch	spa=Spanish	zho=Mandarin

References

de Angelis, G. & L. Selinker 2001. Interlanguage Transfer and Competing Linguistic Systems in the Multilingual Mind. In J. Cenoz, B. Hufeisen & U. Jessner (eds.) *Cross-Linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Clevedon: Multilingual Matters, 42–58.

Baayen, R. H. 2001. *Word Frequency Distributions*. Dordrecht: Kluwer.

Baayen, R. H. 2009. Corpus Linguistics in Morphology: Morphological Productivity. In A. Lüdeling & M. Kytö (eds.) *Corpus Linguistics. An International Handbook*. Vol. 2. Berlin: Mouton de Gruyter, 899–919.

Baayen, R. H. & R. Lieber 1991. Productivity and English Derivation: A Corpus-Based Study. *Linguistics* 29(5), 801–843.

Barz, I. 1995. Komposita im Großwörterbuch Deutsch als Fremdsprache. In I. Pohl & H. Ehrhardt (eds.) *Wort und Wortschatz. Beiträge zur Lexikographie*. Tübingen: Niemeyer, 13–24.

Bauer, L. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.

Boas, H. C. 2003. *Resultative Constructions in English and German*. Stanford: CSLI Publications.

Bohnet, B. 2010. Very High Accuracy and Fast Dependency Parsing is not a Contradiction. In *Proceedings of the 23rd International Conference on Computational Linguistics*. Beijing, 89–97.

Booij, G. E. 2009. Compounding and Construction Morphology. In R. Lieber & P. Štekauer (eds.) *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 201–216.

Booij, G. E. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Borer, H. 2008. Compounds: The View from Hebrew. In R. Lieber & P. Štekauer (eds.) *The Oxford Handbook of Compounds.* Oxford: Oxford University Press, 491–511.

Bybee, J. L. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4), 711–733.

Bybee, J. L. 2013. Usage-based Theory and Exemplar Representations of Constructions. In T. Hoffmann & G. Trousdale (eds.) *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press, 49–69.

Byrnes, H. 2012. Advanced language proficiency. In S. M. Gass & A. Mackey (eds.), *Handbook of Second Language Acquisition*. New York: Routledge, 506–520.

Croft, W. 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Downing, P. A. 1977. On the Creation and Use of English Compound Nouns. *Language* 53(4), 810–842.

Ellis, N. C. 2002. Frequency Effects in Language Processing: A Review with Implications for Theories of Implicit and Explicit Language Acquisition. *Studies in Second Language Acquisition* 24(2), 143–188.

Ellis, N. C. 2012. Formulaic Language and Second Language Acquisition: Zipf and the Phrasal Teddy Bear. *Annual Review of Applied Linguistics* 32, 17–44.

Ellis, N. C., M. B. O'Donnell & U. Römer 2014. Second Language Verb-Argument Constructions are Sensitive to Form, Function, Frequency, Contingency, and Prototypicality. *Linguistic Approaches to Bilingualism* 4(4), 405–431.

Evert, S. & M. Baroni 2007. zipfR: Word frequency distributions in R. In *Proceedings of the 45th Annual Meeting of the ACL, Posters and Demonstrations Session*. Prague, 29–32.

Fandrych, C. & U. Tallowitz 2008. *Klipp und Klar. Übungsgrammatik Grundstufe Deutsch*. Stuttgart: Klett.

Fandrych, C. & M. Thurmair 1994. Ein Interpretationsmodell für Nominalkomposita: linguistische und didaktische Überlegungen. *Deutsch als Fremdsprache* 31, 34–45.

Fanselow, G. 1981. *Zur Syntax und Semantik der Nominalkomposition. Ein Versuch praktischer Anwendung der Montague-Grammatik auf die Wortbildung im Deutschen*. Tübingen: Niemeyer.

Fillmore, C. J. 1988. The Mechanisms of 'Construction Grammar'. *Proceedings of the Berkeley Linguistics Society* 14, 35–55.

Fischer, K. & A. Stefanowitsch 2006. Konstruktionsgrammatik: Ein Überblick. In K. Fischer & A. Stefanowitsch (eds.) *Konstruktionsgrammatik: Von der Anwendung zur Theorie*. Tübingen: Stauffenburg, 3–17.

Fuente, M. J. de la 2002. Negotiation and Oral Acquisition of Spanish L2 Vocabulary: The Roles of Input and Output in the Receptive and Productive Acquisition of Words. *Studies in Second Language Acquisition* 24, 81–112.

Gaeta, L. & D. Ricca 2006. Productivity in Italian Word Formation: A Variable-Corpus Approach. *Linguistics* 44(1), 57–89.

Gaeta, L. & A. Zeldes 2012. Deutsche Komposita zwischen Syntax und Morphologie: Ein korpusbasierter Ansatz. In L. Gaeta & B. Schlücker (eds.) *Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte*. Berlin: De Gruyter, 197–217.

Gawron, J.-M. 2011. Frame Semantics. In C. Maienborn, K. Heusinger & P. Portner (eds.) *Semantics. An International Handbook of Natural Language Meaning*. V. 1. Berlin and Boston: De Gruyter Mouton, 664–687.

Goldberg, A. E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldschneider, J. & DeKeyser, R. M. 2001. Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning* 51, 1–50.

Granger, S. 2002. A Bird's-Eye View of Learner Corpus Research. In S. Granger, J. Hung & S. Petch-Tyson (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam and Philadelphia: John Benjamins, 3–33.

Granger, S., E. Dagneaux, F. Meunier & M. Paquot 2009. *The International Corpus of Learner English. Version 2. Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain.

Hartkamp, S. & K. Schneider-Wiejowski 2010. Die Regelkonformität in der Distribution von Fugenelementen. Eine korpuslinguistische Untersuchung neologistischer Substantivkomposita mit nominalem Erstglied. *Muttersprache* 3/2010, 198–213.

Heidermann, W. 1997. *Grammatiktraining Grundstufe*. Ismaning: Verlag für Deutsch.

Hirschmann, H., A. Lüdeling, I. Rehbein, M. Reznicek & A. Zeldes 2013. Underuse of Syntactic Categories in Falko - A Case Study on Modification. In S. Granger & F. Meunier (eds.) *20 Years of Learner Corpus Research. Looking Back, Moving Ahead*. Louvain: Presses Universitaires de Louvain, 223–234.

Jong, N. H. de, L. B. Feldman, R. Schreuder, M. Pastizzo & R. H. Baayen 2002. The Processing and Representation of Dutch and English Compounds: Peripheral Morphological and Central Orthographic Effects. *Brain and Language* 81(1-3), 555–567.

Klein, W. & C. Perdue 1997. The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research* 13, 301–347.

Klein-Braley, C. & U. Raatz 1982. Der C-Test: Ein neuer Ansatz zur Messung von allgemeiner Sprachbeherrschung. *AKS Rundbrief* 4, 23–37.

Krott, A., R. Schreuder, R. H. Baayen & W. U. Dressler 2007. Analogical Effects on Linking Elements in German Compound Words. *Language and Cognitive Processes* 22, 25–57.

Krummes, C. & A. Ensslin 2014. What's Hard in German? WHiG: A British Learner Corpus of German. *Corpora* 9(2), 191–205.

Kürschner, S. 2005. Verfugung-s-nutzung kontrastiv: Zur Funktion der Fugenelemente im Deutschen und Dänischen. *TijdSchrift voor Skandinavistiek* 26(2), 101–125.

Langacker, R. W. 1987. *Foundations of Cognitive Grammar*. Vol. 1: Theoretical Prerequisites. Stanford: Stanford University Press.

Levi, J. N. 1978. *The Syntax and Semantics of Complex Nominals*. New York: Academic Press.

Lu, X. 2010. Automatic analysis of Syntactic Complexity. *International Journal of Corpus Linguistics* 15(4), 474–496.

Lüdeling, A. 2008. Mehrdeutigkeiten und Kategorisierung: Probleme bei der Annotation von Lernerkorpora. In M. Walter & P. Grommes (eds.) *Fortgeschrittene Lernervarietäten*. Tübingen: Niemeyer, 119–140.

Lüdeling, A. & M. Walter 2010. Korpuslinguistik. In H.-J. Krumm, C. Fandrych, B. Hufeisen & C. Riemer (eds.) *Handbuch Deutsch als Fremd- und Zweitsprache*. Berlin and New York: Walter de Gruyter, 315–322.

MacWhinney, B. (2012). The logic of the unified model. In S. M. Gass & A. Mackey (eds.), *Handbook of Second Language Acquisition*. New York: Routledge, 211–227.

Michel, S. 2009. Schaden-0-ersatz vs. Schaden-s-ersatz. Ein Erklärungsansatz synchroner Schwankungsfälle bei der Fugenbildung von N+N-Komposita. *Deutsche Sprache* 37, 334–351.

Milton, J. 2009. *Measuring Second Language Vocabulary Acquisition*. Clevedon: Multilingual Matters.

Motsch, W. 2004. *Deutsche Wortbildung in Grundzügen*. 2. überarbeitete Auflage. Berlin and New York: Walter de Gruyter.

Müller, F. H. 2004. *Stylebook for the Tübingen Partially Parsed Corpus of Written German (TüPP-D/Z)*. Universität Tübingen, Seminar für Sprachwissenschaft, Technical Report.

Nation, P. 1990. *Teaching and Learning Vocabulary*. Rowley, MA: Newbury House.

Neef, M. 2009. IE, Germanic: German. In R. Lieber & P. Štekauer (eds.) *The Oxford Handbook of Compounding*. Oxford: Oxford University Press, 386–399.

Nesselhauf, N. 2003. The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics* 24(2), 223–242.

Norris, J. M. 2004. *Validity Evaluation in Language Assessment*. PhD Thesis, University of Hawai'i.

Norris, J. M. & L. Ortega 2009. Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics* 30(4), 555–578.

Nübling, D. & R. Szczepaniak 2011. Markmal(s?)analyse, Seminar(s?)arbeit und Essen(s?)ausgabe: Zweifelsfälle der Verfugung als Indikatoren für Sprachwandel. *Zeitschrift für Sprachwissenschaft* 30(1), 45–73.

Nübling, D. & R. Szczepaniak 2013. Linking elements in German Origin, Change, Functionalization. *Morphology* 23, 67–89.

Rainer, F. 1988. Towards a theory of blocking: the case of Italian and German quality nouns. In G. E. Booij & J. van Marle (eds.) *Yearbook of Morphology 1988*. Dordrecht: Kluwer, 155–185.

Rak, Z. 1990. *Substantivische Komposita im Deutschen - semantische Analyse und Beschreibung für die fachbezogene Fremdsprachenausbildung an polnischen Hochschulen*. PhD Thesis, Technische Universität Dresden.

Reznicek, M., A. Lüdeling & H. Hirschmann 2013. Competing Target Hypotheses in the Falko Corpus: A Flexible Multi-Layer Corpus Architecture. In A. Díaz-Negrillo, N. Ballier & P. Thompson (eds.) *Automatic Treatment and Analysis of Learner Corpus Data*. Amsterdam: John Benjamins, 101–124.

Reznicek, M., A. Lüdeling, C. Krummes, F. Schwantuschke, M. Walter, K. Schmidt, H. Hirschmann & T. Andreas 2012. *Das Falko-Handbuch. Korpusaufbau und Annotationen*. Humboldt-Universität zu Berlin, Technical Report Version 2.01, Berlin.

Ringbom, H. 1998. Vocabulary Frequencies in Advanced Learner English: A Cross-linguistic Approach. In S. Granger (ed.) *Learner English on Computer*. London: Addison Wesley Longman, 41–52.

Schiller, A., S. Teufel, C. Stöckert & C. Thielen 1999. *Guidelines für das Tagging deutscher Textcorpora mit STTS*. Universität Stuttgart, Institut für maschinelle Sprachverarbeitung and Universität Tübingen, Seminar für Sprachwissenschaft, Technical Report.

Schlücker, B. 2012. Die deutsche Kompositionsfreudigkeit. Übersicht und Einführung. In L. Gaeta & B. Schlücker (eds.) *Das deutsche als kompositionsfreudige Sprache. Strukturelle Eigenschaften und systembezogene Aspekte*. Berlin: De Gruyter, 1–25.

Schmid, H. 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the Conference on New Methods in Language Processing*. Manchester, UK, 44–49.

Schmid, H., A. Fitschen & U. Heid 2004. SMOR: A German Computational Morphology Covering Derivation, Composition, and Inflection. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*. Lisbon, Portugal, 1263–1266.

Trynkler, A. 2010. Deutsche Komposita und ihre Erlernungsmöglichkeiten in der Oberschule. *Studia niemcoznawcze. Studien zur Deutschkunde* 44, 467–475.

Vyatkina, N., Hirschmann, H. & Golcher, F. (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing* 29, 28–50.

Zeldes, A. 2012. *Productivity in Argument Selection. From Morphology to Syntax*. Berlin and Boston: De Gruyter.

Zeldes, A. 2013. Komposition als Konstruktionsnetzwerk im fortgeschrittenen L2-Deutsch. *Zeitschrift für germanistische Linguistik* 41(2), 240–276.