

A typological comparison of infixes and circumfixes

Tim Zingler
University of Innsbruck
tim.zingler@uibk.ac.at

1. Introduction

While the existence of infixes and circumfixes is acknowledged in many relevant textbooks and handbooks on morphology and/or typology (e.g., Whaley 1997: 117; Hall 2000: 540; Moravcsik 2000: 548), very little is known about their behavior and distribution (cf. also Harris 2010: 100, n. 1). If anything, both types are usually claimed to be cross-linguistically rare, even though no robust evidence in support of this is typically presented. One of the few language families known to have both affix types is Austronesian, and to the extent that the relevant literature provides examples of either type, they tend to come from this family. This situation suggests that a considerable amount of empirical groundwork is required in order to properly understand these two phenomena and to adequately account for them in theories of morphology. The present contribution is intended as a first step toward that goal. It will aggregate the results of two separate studies on the cross-linguistic distribution of infixes and circumfixes, Zingler (2022a, b). Both works are of an inductive nature and as such mostly point to areas that call for further research or that would appear to be of particular theoretical interest. These domains will be highlighted throughout this work.

In Section 2, I will outline some basic facts about infixation and circumfixation and some of the issues and motivations that guided the research underlying this paper. In Section 3, I will juxtapose the main results of the two studies, while Section 4 will be an attempt to find plausible explanations for those patterns. Section 5 is the conclusion, which will sum up the major insights and sketch some crucial desiderata for future work.

2. Infixation and circumfixation: Some basic facts and issues

It should be emphasized at the outset that our relatively limited understanding of infixation only concerns its morphological aspects. That is, the phonological properties of infixes have been studied in cross-linguistic detail and have subsequently been integrated into linguistic theory, particularly Optimality Theory (cf. Yu 2007; Inkelas 2014). Some of the major questions such phonologically oriented works are interested in are which positions of a root an infix can appear in (cf. also Wilson 2014) and under which structural conditions certain infixes might occur outside the root (i.e., as prefixes or suffixes). However, such approaches do not address many other questions that any typologically informed theory of morphology or linguistics will eventually need to answer. First and foremost, phonological accounts obviously pay little attention to the functions that infixes express, and hence the range of meanings that infixes encode has hardly been explored. Put differently, there is much to be gained from a “semasiological” approach to infixation, which essentially takes infixes as the “independent variable” and then explores what kind of functions (i.e., “dependent variables”) this formal type is associated with.

A second major concern is *where* infixes are found, which refers to both their genealogical and their geographical distribution. While Austronesian and, to a lesser extent, Austroasiatic are standardly claimed to be centers of infixation (e.g., Ultan 1975: 172; Naumann & Vogel 2000: 941), there is considerably less agreement concerning their occurrence in the Americas. Specifically, Moravcsik (2000: 548), Akmajian et al. (2001: 19), and Crystal (2008: 243) argue that infixes are relatively widespread in that part of the world, whereas Mithun (1999: 39) explicitly states that they are not common in North America. The overall picture is further complicated by the fact that comparable works on South American languages (e.g., Campbell & Grondona 2012; O'Connor & Muysken 2014) do not make any general claims about the status of infixation. This state of affairs was the impetus for Zingler (2022a), which provides an overview of infixes in American languages on the basis of a convenience sample of about 170 languages. That study primarily relied on reference grammars, with handbooks, textbooks, and journal articles as supplementary sources. The Americas were chosen because the above-mentioned references suggested that the phenomenon is sufficiently common in those two macro-areas to support an in-depth study. Put differently, no other areas with potentially sizable quantities of unknown infixes emerged during the literature review. As a consequence, any claims about the nature and distribution of infixes in this work will be based on that particular study of American infixes, unless explicitly indicated otherwise.

The situation regarding circumfixes is rather different. This structural type has not only failed to attract attention from morphology but has also held little interest for phonology. Hence, circumfixes are even less well explored than infixes, and this crucially impacted the empirical focus of Zingler (2022b). Specifically, the latter work used a worldwide convenience sample of roughly 450 languages to arrive at a first impression of the distribution and behavior of circumfixes. Here, too, I primarily relied on reference grammars and made use of handbooks, textbooks, and journal articles as supplementary sources. Beyond that, the goals of Zingler (2022b) were essentially the same as the ones in Zingler (2022a).

That said, Zingler (2022b) did impose *some* restrictions on the database. One major caveat concerning its sample is that no attempt was made to include a representative number of Austronesian circumfixes. This was due to the fact that this family is extraordinarily large and would thus have had an outsized impact on the database if no such safeguard had been in place. Also, since the goal of typology is to explore cross-linguistic patterns that emerge irrespective of common descent, such an extensive coverage of Austronesian would have undermined the overall purpose of the investigation. Therefore, Austronesian circumfixes were only considered when they expressed a function that was also encoded by non-Austronesian circumfixes, in which case they are part of a larger and independent pattern.

The second major exception to the otherwise inclusive sampling in Zingler (2022b) concerns the Indo-European family. As the discussions in the volumes edited by Müller et al. (2015, 2016) show, many of the Indo-European items that have been or could be described as circumfixes are rather controversial as far as their structural make-up is concerned. As such, an inclusion of (potential) Indo-European circumfixes would often have required a detailed discussion of formal and semantic nuances, which would have gone beyond the goals of Zingler (2022b). The solution to this problem was to entirely exclude Indo-European languages from the sample. This presumably makes Zingler (2022b) one of the few studies with a worldwide sample in which Indo-European languages are underdocumented. Of course, the Indo-European circumfixes are controversial because the relevant languages are well-known, including the analytical challenges they present. In other words, it is likely that many of the non-Indo-European circumfixes in Zingler (2022b) would be seen as equally controversial if the relevant languages were sufficiently understood. As such, that study points toward certain descriptive issues that still require clarification, either on a language-specific or on a general basis.

A perusal of the relevant literature reveals that there are no universally applied definitions of infixation or circumfixation. It is therefore necessary to describe the structural parameters on which the collection of the two databases relied. I will briefly summarize the main criteria here. Details, references, and illustrations can be found in Zingler (2022a: 172-178) and Zingler (2022b: 58-65). With regard to infixes, the major distributional requirement was that they occur *intramorphemically* rather than *intermorphemically*. There is a terminological tradition in which non-peripheral affixes are described as infixes such that the *en-* in English *disenfranchise* would count as an infix. Yet, this is of course very different from the way in which infixation is usually understood. On that more common definition, which also underlies Zingler (2022a), infixes have to split another morph in two. Furthermore, infixes – like all morphs – are linguistic signs, and every item claimed to be an infix therefore has to express a discernible function. This criterion primarily serves to distinguish infixation from epenthesis.

Meanwhile, circumfixes must not be the compositional combination of a prefix and a suffix such that both the preposed and the postposed element independently exist *and* the putative circumfix expresses the combined function of these preposed and postposed items. Like most other criteria mentioned here, this one is often difficult to verify due to the limited information in the sources, but it has to be a general principle. In addition, there are also criteria that apply to both infixes and circumfixes. For instance, in order to be included in the relevant databases, both types have to be instantiated by clearly additive material. That is, phenomena such as root change (including ablaut) do not count as infixation because the allegedly infixed word form is not segmentally longer than the corresponding form without the alleged infix. Similarly, processes in which, for instance, consonant gradation of the word onset expresses a specific function in conjunction with suffixation are not instances of circumfixation because there are no segmental additions at the beginning of the word form in such cases.

Finally, and perhaps most importantly, infixes and circumfixes are affixes and therefore have to be bound to a single word class. This criterion divides into two different issues. On the one hand, it is often impossible to determine on the basis of the evidence available in the sources whether an element really is limited to a single word class (provided that there is even a clear idea of how to define word classes). Put differently, it is somewhat probable that there are more endoclititics and even circumclitics in the languages of the world than currently recognized. On a more general level, however, it can be stated with confidence that both infixes and circumfixes almost exclusively occur with verbs and, to a much lesser extent, with nouns (or with lexemes expressing verbal and nominal concepts). On the other hand, it is unclear how to define “bound” (cf. Haspelmath 2021). In Zingler (2022a, b), this concept was primarily invoked to exclude cases in which full words may be inserted into other words, as in the famous case of English “expletive insertion” (cf. McCarthy 1982). However, this English process seems to be virtually unique, and there were no analogous instances relevant to the classification of circumfixes either. Therefore, the criterion of “boundness” was presumably satisfied to a greater extent across the two studies than any of the other parameters sketched here.

Examples (1), from Mískito, and (2), from Georgian, illustrate items that meet all the above-mentioned criteria for infixes and circumfixes, respectively, to the extent that the relevant sources permit such a conclusion. Both will be discussed below.

- (1) *na<m>pa*
tooth<2SG>
'your tooth'
(Lehmann & Moravcsik 2000: 750)

- (2) *me-cxvar-e*
 AGT₁-sheep-AGT₂
 ‘shepherd’
 (Hewitt 1995: 103)

With respect to (1), the concept of tooth is highly likely to be monomorphemic, which in turn renders it likely that the possessive marker is inserted intramorphemically. Since this marker is assigned the straightforward function of second person possessive, it also qualifies as a full-fledged sign. Furthermore, its sub-syllabic structure suggests that it is indeed a bound unit rather than a potentially free word. Finally, there is no indication in the source that the *m* might be understood as nasalization. With regard to (2), it is unclear how the meaning of agent(ive) could be compositional, and this makes it likely that the element is indeed a single affix, as described in the source. Given this meaning, the item is also a linguistic sign, and there is again nothing to suggest that the preposed and postposed parts are anything but segmental additions. While possessive marking on nouns and argument indexation on verbs often use formally identical items (Siewierska 1998), the two functions are of course distinct, which is why the element in (1) does not violate the principle of attaching to only a single word class. Similarly, the formal structure that instantiates the agentive circumfix in (2) also derives ordinal from cardinal numerals. But since these functions are unrelated, the two structures are not the same sign. Therefore, (2) does not illustrate a violation of word class loyalty either.

3. Data

Zingler (2022a) found 61 infixes from 48 American languages, which belong to 32 families. Meanwhile, Zingler (2022b) found 83 circumfixes from 59 languages, which belong to 33 families. (The genealogical classification is based on *Glottolog*; Hammarström et al. 2022. Each isolate counts as a separate family.) Tables 1 and 2 show which functions the infixes and the circumfixes express, respectively. The classification of the functions is further described below, whereas the potential explanatory value that the different functions have for each of the affix types will be a major focus of the next section.

Table 1: The functions of infixes.

<i>Function</i>	<i>Number of languages</i>
Number (including collective, distributive, pluractional)	18
Tense/aspect/mood (modality)/evidentiality (TAME)	14
Person	9
Voice/valency	8
Miscellaneous	12

Table 2: The functions of circumfixes.

<i>Function</i>	<i>Number of languages</i>
Negation (including privative, irrealis, prohibitive)	20
Nominalization	8
Mood/modality	7
Miscellaneous	48

The two tables reveal several obvious asymmetries. First and foremost, the infixes are a much more semantically coherent set than the circumfixes. The former largely divide into four semantic domains, with the “miscellaneous” category accounting for less than 20% of the data. Yet, even this is deceptive because several of the infixes subsumed under the “miscellaneous”

label express some kind of intensification. Hence, intensification could be regarded as a fifth major function expressed by the infixes, which would further reduce the proportion of infixes with “miscellaneous” functions.

By contrast, the circumfixes are a very heterogeneous class in which the “miscellaneous” group is by far the largest and accounts for more than half the data. In fact, it would be futile to list which functions are included among the “miscellaneous” circumfixes because they subsume most major inflectional and derivational meanings. The main exceptions are from the nominal domain: demonstratives and definiteness markers, as well as noun class and gender markers. The absence of these categories from the database is also essentially the only semantic property that the circumfixes share with the infixes. That is, among the categories that they do express, there is essentially no overlap. The only candidate might be mood/modality, but only a few of the TAME infixes actually express mood and/or modality. Furthermore, the concept of mood/modality is so wide (not to say “ill-defined”) that this degree of convergence appears unremarkable on the surface.

In addition to their semantic patterns, the two affix types also show interesting geographical distributions. While these will not be discussed any further here, they should be mentioned in the interest of completeness. In discussing areal patterns, I will again follow the classification in *Glottolog*. My circumfix sample only contains two items from the macro-area of Australia, which is a conspicuously low number compared to all the other macro-areas. By contrast, the macro-area of Papunesia accounts for 25 of the circumfixes, and circumfixation is relatively common in that part of the world even if Austronesian is factored out. On the other hand, there are no infixes in most of the northernmost families of North America (specifically, Eskimo-Aleut, Iroquoian, and Athabaskan-Eyak-Tlingit) but noticeably many across Mesoamerica.

Of course, the results obtained here may simply be artifacts of the convenience samples on which they rest. However, both studies appear sufficiently comprehensive to suggest that there are indeed principled differences between the distributions of the two types. Crucially, there is nothing to suggest that the two databases are incommensurate because one is based on a worldwide sample and the other one is only based on American languages. (But of course, the baseline expectation is that Austronesian infixes would diversify the range of infixal functions.) The next section will propose certain explanations for the most prominent associations between form and function described above.

4. Analysis

The basic explanations that I will pursue for the patterns in Tables 1 and 2 are of the diachronic kind. Of course, first-hand historical documentation is hardly available for any of the languages in either of the samples, and I will therefore rely on the diachronic pathways established for infixes and circumfixes more generally. In addition, I will supplement these accounts with insights gained from cross-linguistic studies of grammaticalization (or “constructionalization”). Put differently, I will assume that the synchronic functions of infixes and circumfixes are an indicator of their diachronic development and that, more broadly, diachrony can help to make sense of synchronic phenomena (Bybee 2015).

Since the infixes constitute a fairly homogeneous group, explanations for their functional associations are presumably more straightforward than they are in the case of the highly diverse circumfixes. In fact, Yu (2007: ch. 5) provides a detailed account of the historical developments by which infixes arise. He describes three major paths: reduplication mutation, metathesis, and entrapment. The concept of reduplication mutation roughly refers to a process during which an erstwhile peripheral (i.e., “adfixal”) reduplicant is reanalyzed as part of the root, and a part of the base is simultaneously reanalyzed as the reduplicant. As such, the reduplicant “accidentally” lands inside the root. Schematically, this can be thought of as $ab-ABCD \rightarrow AB<ab>CD$, where

lower-case letters mark affixal segments and upper-case letters mark root segments. While Yu (2007: ch. 5) adds many more complexities to this account, which cannot be verified for the sample of Zingler (2022a), what matters for the present semasiological approach is simply that reduplication as such is often the origin of infixes. This is crucial in that the meanings expressed by reduplication are also widely found among the American infixes. In particular, this is true of the iconic functions of reduplication, where a repetition of form expresses a repetition of meaning (cf. Moravcsik 1978: 316-325). Ultimately, the vast majority of number-marking infixes fall into this category, as do most of the TAME infixes since these largely express imperfective aspects such as repetitive, iterative, etc. In addition, the intensifiers that account for several of the “miscellaneous” infixes are also iconic and are thus also plausibly derived from reduplicative constructions. As I discuss in Zingler (2022a: 209-212), many of the infixes expressing these functions are indeed (partly) reduplicative and, equally crucially, none of the other infixal functions are expressed by reduplicative infixes. As such, a major portion of the infix database can straightforwardly be attributed to an iconic origin that leads to infixation via reduplication (mutation).

The process of metathesis is simpler and better-known than reduplication mutation, but also harder to detect in synchronic data. In Zingler (2022a), I argued that the infixes most likely to have emerged by this process are the ones that express voice/valency. This claim is based on Bybee (1985), who shows that functions that are most relevant to the meaning of the verb stem also occur close to the verb stem (which is another instance of iconicity). She argues explicitly that voice and valency are the most relevant categories in the verbal domain (Bybee 1985: 4-5), from which it should follow that voice/valency affixes tend to be immediately adjacent to the stem. While she does not investigate this particular prediction in detail, her claims are borne out for all the categories that she does explore. Other amenable generalizations in this context are that metathesis typically affects adjacent and short units and that affixes are typically short. Hence, voice/valency affixes should meet both the structural and the distributional prerequisites for undergoing metathesis and for thus ending up inside the neighboring root. Of course, the factors that give rise to metathesis are primarily phonological, but these cannot be covered here.

The last process that Yu (2007) discusses is “entrapment.” This refers to another kind of reanalysis in which an erstwhile adfix occurs between two elements that ultimately cumulate to become a single morph such that the former adfix develops into an infix due to the semantic changes taking place around it. Yu (2007: 148) limits this process to cases in which one surrounding morph is an affix and the other one is a root, but there do not seem to be compelling reasons to exclude scenarios in which both surrounding items are stems. For a prefixal origin, this could thus be schematically represented as *Stem 1* – [*Prefix* – *Stem 2*] → [*Stem 1* – *Infix* – *Stem 1*]. It follows that entrapment is a process that may take place when affixes are not “externalized” (Haspelmath 1993). As before, the sparse synchronic data do not permit specific entrapment analyses for any of the infixes. However, Bybee’s (1985) account is yet again suggestive at a more general level. That is, she argues that person markers are least relevant to the meaning of a verb and are thus also furthest removed from the verb stem and generally less tightly fused to it. It seems to be a defensible assumption that a lower degree of formal bonding with the stem is a necessary condition for dissociating from that stem and for assuming a more neutral position between the two parts of the new stem. If so, person markers would be the most likely category of the ones in Table 1 to develop infixal exponents via entrapment. In that case, all the major functional domains in Table 1 would be accounted for by one of Yu’s (2007) diachronic scenarios.

While this account of the functional patterns shown by infixes is perforce speculative, it should also be mentioned that there are a few items in the database that are not captured by any of the diachronic trajectories (for instance, a stem formation marker in Central Sierra Miwok).

Since the circumfixes show a much wider range of functions than the infixes, they would also seem to require a much more expansive explanation. Yet, such an explanation is complicated by the lack of previous work on the topic. Therefore, the following remarks are based on more general principles of language change and will have to be supplemented or replaced by more detailed diachronic work. Two of the very few works that make substantive claims about (the emergence of) circumfixes are Greenberg (1980) and Harris (2010). Both are essentially in agreement that circumfixes arise from the reanalysis of an originally independent prefix and an originally independent suffix as a single affix. There is nothing in my data that would argue against this scenario, and this account has the additional merit that it would explain why circumfixes have a much wider range of functions than infixes to begin with. Simply put, there are essentially no semantic constraints on prefixes or suffixes, and hence there also should not be any semantic constraints on circumfixes that arise from prefixes and suffixes.

That said, there is an obvious link between circumfixation and negation. Of the 20 negation markers in the database, 13 express “standard negation” in the sense of Miestamo (2005), whereas the remaining seven items express privative, irrealis, or prohibitive functions. Hence, the latter primarily differ from standard negation in that they do not scope over a declarative predicate, but the decision to subsume irrealis under negation is more controversial (cf. Zingler 2022b: 71 for discussion). While I do not know of any concrete evidence for how negation circumfixes arise in a language, certain diachronic patterns nevertheless permit a relatively plausible explanation. The crucial ingredient of such an account is arguably the negative cycle, during which a negator loses its pragmatic force and comes to be supported by a second negator, which eventually becomes the main marker of negation. This process is well-documented for French but is also observable in many other languages (van Gelderen 2008). Presumably, then, if the two parts of the negative construction are found on opposite sides of the clausal head, and the stage at which they co-occur is sufficiently long for both to morphologize, a negation circumfix may emerge. The fact that negation circumfixes are rare overall therefore seems to confirm the intuition that these diachronic developments, or at least their combination, are also rare. In this context, it is also important to highlight that negators show a tendency to be preposed. As such, they meet a necessary requirement for the emergence of circumfixes that the exponents of most other categories satisfy to a much lesser degree. Bybee (1985: 177) even finds a slight prefixation preference among the negation affixes in her worldwide sample. It is this prefixing bias that gives negators a head start in the “circumfix race” compared to the markers of virtually any other category. (For a functional explanation as to why negators tend to occur in preposed position, cf. Dryer 1988: 102; Berg 2020: 386.)

The second major set among the circumfixes consists of nominalizers. This group is arguably in even greater need of explanation given the cross-linguistic fact that nouns show considerably fewer prefixes than verbs (Cutler et al. 1985: 730; Seifart et al. 2018: 5723). Hence, nouns violate a crucial requirement for the development of circumfixes. Yet, while any diachronic explanation of this pattern will also have to be speculative, one angle seems somewhat promising. In my sample, most of the stems that form the input for nominalization via circumfixes are verbal. The unexpected clustering of circumfixes in the domain of nominalization might thus find an explanation if it could be shown that the nominalization circumfixes originally encoded verbal categories and only became nominalizers via semantic change. There is some tentative evidence for that scenario. While Cristofaro (2019) shows that nominalizers mostly derive from nouns with general meanings such as ‘person,’ ‘thing,’ or ‘place,’ the grammaticalization literature also reveals that nominalizers seem to emerge in the context of participles (cf. Kuteva et al. 2019: 334, 434-4). Hence, if a single verbal stem is separately marked for both nominalization and non-finite status, there are two items that could eventually be reanalyzed as a single nominalizer. That these categories are likely to co-occur also follows from their semantic compatibility, specifically, from the fact that participles and

nominalizers assume intermediate positions on the noun-verb cline (cf. Hopper & Thompson 1984: 704; Comrie & Thompson 2007: 346). Crucially, this account leaves open the possibility that the prefixal part of the eventual circumfixes started out as a verbal marker (i.e., non-finiteness), which would be the unmarked case. As before, the fact that this scenario appears highly specific and thus unlikely is not a problem because the outcome of that scenario is indeed very infrequent.

There is also another conceivable diachronic path for nominalization circumfixes, though support for this claim comes from only one family in my data. The Tacanan branch of the Pano-Tacanan family has several different nominalizing circumfixes with *e-* as the first part and a variety of items as the second part (cf. Guillaume 2008: 435 for Cavineña; Vuillermet 2012: 313 for Ese Ejja). Crucially, the initial *e-* is homophonous with a former noun classifier that has since lost its function. It therefore stands to reason that this classifier was only “accidentally” present in the nouns originally formed by nominalizing suffixes and that it was absorbed by these suffixes because it was no longer associated with a function. On that reading, then, these circumfixes arose from the combination of a suffix and a meaningless element.

Any explanation of the remaining categories will have to be even less comprehensive. For instance, Bybee et al. (1990: 29-34) show that aspect markers are much more frequently proposed than tense markers. That arrangement might pave the way for a process in which a perfective prefix and a past tense suffix coalesce into a single circumfix, given their functional near-equivalence. However, the TAME category most widely found among the circumfixes is mood/modality rather than tense/aspect, and no analogous developments suggest themselves in the former domain. Among the other functional clusters in the data is a small set of adjectival degree markers. Greenberg (1966 [2005]: 40) suggests that the Hungarian superlative circumfix arose via the further affixation of the comparative form. Such a development might be further facilitated if the comparative and superlative function are only weakly distinguished (cf. Hewitt 1995: 49; Cuzzolin & Lehmann 2004: 1215 for Kartvelian), in which case the second affix might primarily provide clarification.

Overall, then, the emergence of circumfixation is still largely a mystery, but some advances toward an understanding of this phenomenon can be made once independently established cross-linguistic phenomena such as the negative cycle or the suffixing preference are taken into account. Further progress on this topic will mostly depend on the work of specialists on the relevant languages.

5. Conclusion

Among the ideas espoused here is the claim that the functional patterns of the two affixation strategies can usually be explained with reference to their diachrony. Another conclusion is that infixation seems to be more widely distributed than circumfixation. This claim is based on the fact that I have data from roughly the same number of language families for both types, even though the infixes are only from the Americas. Hence, infixes seem to arise more easily than circumfixes, which would be plausible in that the emergence of infixes almost exclusively depends on formal mechanisms, whereas both formal and functional criteria have to be met in order for circumfixes to arise. On the other hand, there are several language families that make rather extensive use of circumfixes (Chukotko-Kamchatkan, Kartvelian, Austronesian), while infixes seem to be highly sporadic in every American language (family) they occur in. Hence, if infixes are indeed a “wider” property, circumfixes in turn might be a “deeper” feature and might even have to be reconstructed for some proto-languages.

Above all else, this work has shown that infixes and circumfixes are not “rare” in any absolute sense. There are dozens of language families that have at least one of the two types and many that have both. Instead, the relative infrequency of these types compared to prefixes

and suffixes is straightforwardly tied to their diachronic emergence (cf. Harris 2008, 2010). In conjunction, then, both the synchronic and the diachronic properties of infixes and circumfixes call for much further work. Put more bluntly, our current state of knowledge regarding these two types is somewhat of an embarrassment for morphological theory.

Desiderata for future research are as obvious as they are variegated. Each of the diachronic explanations for the circumfixes suggested here would ideally be supported by concrete evidence from at least a few languages. With regard to both affix types, synchronic explanations would also have to be pursued, especially in light of psycholinguistic arguments of why infixes *should* be rare (Cutler et al. 1985: 751-752). In addition to the obvious need for more data (especially from Austronesian languages), there are also many analytical issues that will require meticulous research. For instance, several of the circumfixes in my database appear to be at least partly compositional, whereas several of the infixes might actually be glottalization features rather than segmental additions. As described in Section 2, these issues are essential for the identification of the relevant types, and my databases and those similar to mine might have to be considerably revised once the relevant semantic and phonetic issues are properly understood. In addition, more comprehensive morphosyntactic analyses will have to determine to what extent the items considered infixes and circumfixes here can occur with different word classes and might thus be better conceived of as endoclitics and circumclitics, respectively.

References

- Akmajian, Adrian, Richard Demers, Ann Farmer & Robert Harnish. 2001. *Linguistics*. 5th edn. Cambridge, MA: MIT Press.
- Berg, Thomas. 2020. Ordering biases in cross-linguistic perspective: The interaction of serial order and structural level. *Linguistic Typology* 24(2). 353-397.
- Booij, Geert, Christian Lehmann & Joachim Mugdan (eds.). 2000. *Morphology*, vol. 1. Berlin: De Gruyter.
- Bybee, Joan. 1985. *Morphology*. Amsterdam: Benjamins.
- Bybee, Joan. 2015. *Language change*. Cambridge: Cambridge University Press.
- Bybee, Joan, William Pagliuca & Revere Perkins. 1990. On the asymmetries in the affixation of grammatical material. In William Croft, Keith Denning & Suzanne Kemmer (eds.), *Studies in diachronic typology for Joseph H. Greenberg*, 1-42. Amsterdam: Benjamins.
- Campbell, Lyle & Verónica Grondona. 2012. *The indigenous languages of South America*. Berlin: De Gruyter.
- Comrie, Bernard & Sandra Thompson. 2007. Lexical nominalization. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 3, 334-381. Cambridge: Cambridge University Press.
- Cristofaro, Sonia. 2019. Nominalization in cross-linguistic diachronic perspective. In Robert Zariquiey, Masayoshi Shibatani & David Fleck (eds.), *Nominalization in languages of the Americas*, 169-194. Amsterdam: Benjamins.
- Crystal, David. 2008. *A dictionary of linguistics and phonetics*. 6th edn. Malden: Blackwell.
- Cutler, Anne, John Hawkins & Gary Gilligan. 1985. The suffixing preference: A processing explanation. *Linguistics* 23(5). 723-758.
- Cuzzolin, Pierluigi & Lehmann, Christian. 2004. Comparison and gradation. In Geert Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas (eds.), *Morphology*, vol. 2, 1212-1220. Berlin: De Gruyter.
- Dryer, Matthew. 1988. Universals of negative position. In Michael Hammond, Edith Moravcsik & Jessica Wirth (eds.), *Studies in syntactic typology*, 93-124. Amsterdam: Benjamins.
- Gelderen, Elly van. 2008. Negative cycles. *Linguistic Typology* 12(2). 195-243.
- Greenberg, Joseph. 1966 [2005]. *Language universals*. Berlin: De Gruyter.
- Greenberg, Joseph. 1980. Circumfixes and typological change. In Elizabeth Traugott, Rebecca Labrum & Susan Shepherd (eds.), *Papers from the fourth international conference on historical linguistics*, 233-241. Amsterdam: Benjamins.
- Guillaume, Antoine. 2008. *A grammar of Cavineña*. Berlin: De Gruyter.
- Hall, Christopher. 2000. Prefixation, suffixation and circumfixation. In Booij et al. (eds.), 535-545. Berlin: De Gruyter.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2022. *Glottolog 4.6*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.6578297>. Available online at <http://glottolog.org> (last access 10-25-2022).

- Harris, Alice. 2008. Explaining exuberant agreement. In Þórhallur Eypórssón (ed.), *Grammatical change and linguistic theory*, 265-283. Amsterdam: Benjamins.
- Harris, Alice. 2010. Explaining typologically unusual structures: The role of probability. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rethinking universals*, 91-103. Berlin: De Gruyter.
- Haspelmath, Martin. 1993. The diachronic externalization of inflection. *Linguistics* 31(2). 279-309.
- Haspelmath, Martin. 2021. Bound forms, welded forms, and affixes: Basic concepts for morphological comparison. *Voprosy Jazykoznanija* 1(1). 7-28.
- Hewitt, B. G. 1995. *Georgian*. Amsterdam: Benjamins.
- Hopper, Paul & Sandra Thompson. 1984. The discourse basis for lexical categories in Universal Grammar. *Language* 60(4). 703-752.
- Inkelas, Sharon. 2014. *The interplay of morphology and phonology*. Oxford: Oxford University Press.
- Kuteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog & Seongha Rhee. 2019. *World lexicon of grammaticalization*. 2nd edn. Cambridge: Cambridge University Press.
- Lehmann, Christian & Edith Moravcsik. 2000. Noun. In Booij et al. (eds.), 732-57.
- McCarthy, John. 1982. Prosodic structure and expletive infixation. *Language* 58(3). 574-590.
- Miestamo, Matti. 2005. *Standard negation*. Berlin: De Gruyter.
- Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge: Cambridge University Press.
- Moravcsik, Edith. 1978. Reduplicative constructions. In Joseph Greenberg, Charles Ferguson, Edith Moravcsik (eds.), *Universals of human language, vol. 3: Word structure*, 297-334. Stanford: Stanford University Press.
- Moravcsik, Edith. 2000. Infixation. In Booij et al. (eds.), 545-552.
- Müller, Peter, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.). 2015. *Word-Formation*. Vols. 1-3. Berlin: De Gruyter.
- Müller, Peter, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.). 2016. *Word-Formation*. Vols. 4-5. Berlin: De Gruyter.
- Naumann, Bernd & Petra Vogel. 2000. Derivation. In Booij et al. (eds.), 929-943.
- O'Connor, Loretta & Pieter Muysken (eds.). 2014. *The native languages of South America*. Cambridge: Cambridge University Press.
- Seifart, Frank, Jan Strunk, Swintha Danielsen, Iren Hartmann, Brigitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich, Nivja de Jong & Balthasar Bickel. 2018. Nouns slow down speech across structurally and culturally diverse languages. *National Academy of the Sciences of the United States of America (PNAS)* 115(22), 5720-5726.
- Siewierska, Anna. 1998. On nominal and verbal person marking. *Linguistic Typology* 2(1). 1-55.
- Ullman, Russell. 1975. Infixes and their origins. In Hansjakob Seiler (ed.), *Linguistic Workshop III*, 157-205. Munich: Fink.
- van Gelderen, Elly (see Gelderen).
- Vuillermet, Marine. 2012. *A grammar of Ese Ejja, a Takanan language of the Bolivian Amazon*. Lyon: Université Lumière Lyon 2 dissertation.
- Whaley, Lindsay. 1997. *An introduction to typology*. Thousand Oaks: SAGE.
- Wilson, Jennifer. 2014. Evidence for infixation after the first syllable: Data from a Papuan language. *Phonology* 31(3). 511-23.
- Yu, Alan. 2007. *A natural history of infixation*. Oxford: Oxford University Press.
- Zingler, Tim. 2022a. Infixation in the Americas: A cross-linguistic survey. *International Journal of American Linguistics* 88(2). 171-228.
- Zingler, Tim. 2022b. Circumfixation: A semasiological approach. *Word Structure* 15(1). 55-113.