UNIVERSITY OF ICELAND

Icelandic and Faroese

A usage-based cognitive analysis of morphological change

Jón Símon Markússon

Dissertation towards the degree of Doctor of Philosophy
2024

SCHOOL OF HUMANITIES
FACULTY OF ICELANDIC AND COMPARATIVE CULTURAL STUDIES

Íslensku- og menningardeild Háskóla Íslands
hefur metið ritgerð þessa hæfa til varnar
við doktorspróf í íslenskri málfræði

Reykjavík, 22. maí 2024

Gauti Kristmannsson
deildarforseti

The Faculty of Icelandic and Comparative Cultural Studies
at the University of Iceland
has declared this dissertation eligible for defence
leading to a Ph.D. degree in Icelandic Linguistics

Doctoral Committee:
Þórhallur Eyþórsson, supervisor
Hjalmar P. Petersen
Katrín Axelsdóttir

Icelandic and Faroese: A usage-based cognitive analysis of morphological change.

This thesis was supported by The Icelandic Research Fund, grant no. 174253-015.

ISBN 978-9935-9736-4-1
ORCID orcid.org/0000-0002-3280-2652

Abstract

The current thesis presents three published articles on inflectional change in Insular Nordic (Icelandic and Faroese). Papers I and II deal with change in Icelandic, while Paper III focuses on Faroese. The three articles are related through employment of the usage-based cognitive approach, which views the structure of grammar as emergent from prior linguistic experience, assuming a central role for language use as the mechanism of linguistic innovation and change. Usage-based cognitive studies typically invoke factors such as frequency and schematicity to account for the varying degrees of productivity that inflectional classes exhibit cross-linguistically. Such studies also make recourse to domain-general cognitive processes like analogy, categorisation, entrenchment, and statistical learning as determinants in the direction of change. Crucially, the usage-based cognitive approach posits rich memory for language. Thus, the cognitive prerequisites for storage and utility of linguistic experience as informative of usage choices are no different from those which inform our interactions with the wider world generally.

Papers I and II account for the limited productivity of the low frequency X*ó*/*æ*T-microclass which, before the addition of borrowed *blók* 'wretch, non-entity', contained five Icelandic feminine nouns in nominative/accusative plural *-ur*: *bók* 'book', *bót* 'patch', *brók* 'trousers', *nót* 'fishing net', and *rót* 'root' only, cf. plural *bækur*, *bætur*, *brækur*, *nætur*, *rætur*. Productivity is equated with the rate at which feminine grammatical gender is assigned to masculine nouns, while the motivation for such treatment is considered to be phonetic coherence with varyingly schematic feminine classes in plural *-ur*. Specifically, Paper I accounts for the different rates at which Icelandic masculine plural forms in with final *-ur* — be that sequence an ending or part of the stem etymologically — undergo reanalysis as feminine. Crucially, around 15% of nouns in Icelandic end in plural *-ur*: almost 92% of these are feminine, while all others are masculine. Further, syncretism in nominative/accusative plural is relatively rare among masculine nouns, but exceptionless among feminines. Also without exception, the relevant forms in plural *-ur* are always syncretic, irrespective of a noun's gender.

Interestingly, in the minority of cases, plurals such as masculine *eigendur* 'owners', *fætur* 'feet', and *vetur* 'winters' alternate with overtly feminine definite forms such as pl.**def**. *eigendur**nar***, *fætur**nar***, *veturnar*, cf. original and more frequent masc.nom.pl.**def**.

*eigendur**nir**, fætur**nir**, vetur**nir***, masc.acc.pl.def. *eigendur**na**, fætur**na**, vetur**na***. Additionally, masculine forms in plural *-ur* sometimes occur with feminine modifiers and determiners. Paper I argues that, given the highly schematic nature of the full set of nouns in plural *-ur*, reanalysis as feminine might be expected at a rate proportionate to the frequency of corresponding masculine forms — all other things being equal. However, based on corpus data for Icelandic, Paper I reports a mismatch in frequency between sets of doublets defined in terms of gender. Through employment of Bybee's network model, with some innovative notational features, Paper I demonstrates that graded phonetic structure of a broader feminine subtype in plural *-ur* — as it centres around the X*ó*/*æ*T-microclass — impacts the rate of reanalysis by means of a gang effect, which is viewed as a function of analogy, i.e. the process by which existing knowledge is extended to new contexts.

In a similar vein, Paper II examines the limited productivity of the Icelandic X*ó*/*æ*T-microclass. In the article, productivity is equated with the occasional inflection of feminine *blók* 'nonentity' and forms of neuter *kók* 'Coke™' according to the morphophonological alternation exhibited by e.g. sg. *b<u>ó</u>k* ~ pl. *b<u>æ</u>k**ur***, *r<u>ó</u>t* ~ *r<u>æ</u>t**ur***, cf. *bl<u>ó</u>k* ~ *bl<u>æ</u>k**ur***, *k<u>ó</u>k* ~ *k<u>æ</u>k**ur***. These new plural forms are taken as evidence for the — albeit highly limited — productivity of the microclass. Indeed, doublet forms in both paradigms also pattern with inflection classes of higher type frequency, cf. pl. *blók**ir***, *kók**ir***, like fem. pl. *mynd**ir*** 'pictures', both of which prove more frequent than plural *blækur* and *kækur*. According to the usage-based cognitive approach to language, the impact of varyingly large and varyingly schematic classes is indeed expected to correlate with graded degrees of productivity.

Some have implied contrastive motivation and, therefore also, distinct cognitive mechanisms for the deduction of plural *blækur* from sg. *blók* and of plural *kækur* from sg. *kók*. Specifically, the opinion has been expressed that plural *kækur* belies "real" language use because the form only occurs in humorous contexts. Plural *blækur*, on the other hand, is considered "real" language use. However, it is clear that both forms are based on the pattern of alternation exemplified by e.g. sg. *bók* ~ pl. *bækur*. Therefore, Paper II seeks to dispel the idea that different motivations for deduction are at play as a misunderstanding of analogy. This objective is achieved through reference to schematicity, semantics, and pragmatics, as well as to Icelandic corpus data. By this means, Paper II demonstrates that new membership in the microclass is graded as a function of limited productivity. Finally, Paper II shows that innovative *kækur* and *blækur* are deduced by identical means, i.e. via analogy.

Paper III deals with analogical change in Faroese, with specific focus on levelling of the intricate vowel alternations which characterised the inflection of Old West Nordic *u*-stems.

Crucially, levelling is defined as the extension of a stem variant to a cell in which it did not occur previously, as opposed to one form "changing into" another. The Faroese descendants of Old West Nordic *u*-stems have undergone significant levelling, as is evident from the paradigm of Far. *vøllur* 'field, grassy ledge on a rock face, (sports) pitch, airport': all cells of the paradigm contain the variant *vøll-* (< OWN nom.sg., acc.sg., acc.pl., dat.pl. *vǫll-*), while some also contain *vall-* (OWN gen.sg., gen.pl. *vall-*). The variant OWN (dat.sg., nom.pl.) *vell-* has been completely eradicated (see Markússon 2022b).

Conversely, the paradigm of Far. *fjørður* 'fjord, inlet/bay, sound/strait' has retained all stem variants, cf. Far. *fjørð-* (< OWN nom.sg., acc.sg., acc.pl., dat.pl. *fjǫrð-*), *firð-* (< OWN dat.sg., nom.pl. *firð-*), *fjarð-* (< OWN gen.sg., gen.pl. *fjarð-*). Further, it has extended those variants to other cells of the paradigm, cf. innovative dat.sg. *fjørði*, which exists beside older *firði*, innovative nom./acc.pl. *fjørðir* and *fjarðir* beside older *firðir*, and innovative dat.pl. *fjarðum* and *firðum*, which live alongside older *fjørðum*.

The basic forms of paradigms, i.e. those from which new inflectional forms take their stem, and the factors that establish them have typically been defined according to either of two opposing theoretical approaches. The first attributes basic status on account of so-called 'markedness', i.e. the perspective that new forms in a paradigm are likely to be based on existing ones that express semantically "natural" and/or "neutral" values, such as *singular* and *nominative*. In other words, such "unmarked" forms serve as basic. The opposing approach posits frequency as the determining factor. Thus, levelling proceeds from the most frequent member(s) of the paradigm, due to a correlation between frequency of use and its impact on the strength of representation in memory. In other words: frequent forms are better represented than less frequent forms and more readily accessible in moments of memory lapse. Therefore, frequent forms are most likely to be used as a base when the "correct" form evades the language user.

Paper III utilises Faroese corpus data in order to demonstrate that the basic forms of Far. *vøllur* and *fjørður* — both of which refer to topographical entities and occur as complex place names — are established on the basis of frequency, rather than semantics. Paper III argues that due to the overall low frequency of forms of Far. *vøllur*, the most frequent stem variant, i.e. *vøll-*, was extended to the whole paradigm, while *vell-* was easily forgotten. Conversely, the high token frequency of dat.sg. *firði* meant that it was well represented in memory and, therefore, easily accessible in moments of temporary memory lapse. Paper III argues that this property of dat.sg. *firði* triggered spread of the stem variant *firð-* to the dative plural through the context [*í*/*á*/*úr* + dat.], where younger dat.pl. *firðum* takes older *fjørðum* over in

frequency. Further, presence of the stem variant *firð-* in all plural cells of the paradigm, cf. also nom./acc.pl. *firðir*, facilitated association of the form *firð-* with the meaning plural. Subsequently, an attempt was made to level the singular portion of the paradigm in favour of the variant *fjørð-* to contrive the formal opposition sg. *fjørð-* : pl. *firð-*. However, the new form never took over the role of older dat.sg. *firði* on account of the high token frequency of the latter.

The current thesis demonstrates that factors such as frequency and schematicity impact choices made in on-line language use as a function of stored experience with language. Further, if the course of language change correlates with the distributional properties of these factors in the acquired grammar, it follows logically that the usage events which incrementally facilitate change reflect the linguistic experience whence the grammar emerges. Moreover, the fact that language change is a function of language use demonstrates that the structure of grammar is an emergent and dynamic system, rather than one whose adaptive properties are constrained by genetic endowment and ontogeny. Therefore, in light of the conclusions drawn in Papers I–III, the current thesis also showcases the applicability of usage-based cognitive theory as a means to account for the direction of morphological change.

Ágrip

Ritgerð þessi fjallar um þrjár rannsóknir, sem hver hefur verið gefin út í sér grein í viðurkenndu tímariti um málvísindi. Greinarnar þrjár eru birtar strax á eftir 7. kafla. Í köflum 1–7 er vísað í greinarnar sem Paper I (1. greinin), Paper II (2. greinin) og Paper III (3. greinin). Í fyrstu greinunum tveimur er hugað að beygingarbreytingum og -nýjungum í íslensku en í þeirri þriðju er einblínt á áþekkt efni í færeysku. Í hverri grein var beitt málnotkunarnálgun á viðkomandi efni en slík nálgun gerir áhrifum almennra hugrænna ferla hátt undir höfði, s.s. útvíkkun hliðstæðrar þekkingar (e. *analogy*), flokkunar (e. *categorisation*), rótfestu (e. *entrenchment*) og hlutfallslegs lærdóms (e. *statistical learning*), á bæði málnotkun og stefnu málbreytinga. Þar að auki skera þættir sem ekki varða málkerfið, s.s. notkunartíðni og formleg líkindi ólíkra eininga, úr um stefnu þessara ferla að töluverðu leyti. Þannig er málkerfið álitið vera ævinlega í mótun á grundvelli fyrri reynslu af málinu. Til þess að málnotkun og málbreytingar endurspegli áhrif umræddra hugrænna þátta er nauðsynlegt að gera ráð fyrir ríkulegu minni fyrir mannlegt mál (e. *rich memory for language*). Þannig séu hugrænar forsendur fyrir geymslu málfræðilegra upplýsinga þær sömu og eru fyrir geymslu annars konar reynslu í minni.

Fyrstu greinarnar tvær fjalla um takmarkaða virkni kvenkyns X*ó*/*æ*T-örflokksins (e. *microclass*), sem hefur lága mynsturstíðni (e. *type frequency*). Á fornískenskum tíma samanstóð flokkurinn af nafnorðunum *bók*, *bót*, *brók*, *glóð*, *nót* og *rót*, sbr. ft. *bækur*, *bætur*, *brækur*, *nætur*, *rætur*. Virkni er ákvörðuð á grundvelli þess hve oft málfræðilegu kvenkyni er úthlutað til upprunalegs karlkynsorðs og er ástæða úthlutunarinnar talin vera misjafnlega mikil hljóðfræðileg líkindi við nafnorð úr ólíkum kvenkynsflokkum með ft. *-ur*. Í þessu sambandi fjallar fyrsta greinin um hversu oft ólíkar karlkynsmyndir með umræddri endingu – – hvort sem sama runa reynist beygingarending eða hluti af stofni — eru umtúlkaðar sem kvenkynsmyndir. Um 15% nafnorða í íslensku fá endinguna ft. *-ur* en næstum því 92% þeirra eru kvenkyns; önnur eru karlkyns. Þar að auki eru samhljóða myndir í nefnifalli og þolfalli fleirtölu hjá karlkynsorðum afar sjaldgæfar, en samsvarandi myndir kvenkynsnafnorða eru undantekningarlaust samhljóða (e. *syncretic*). Enn fremur eru nefnifalls- og þolfallsmyndir með ft. *-ur* afdráttarlaust samhljóða burtséð frá málfræðilegu kyni.

Áhugavert er að karlkynsmyndir eins og ft. *eigendur*, *fætur*, *vetur* víxlast á í beygingu við ákveðnar myndir sem að forminu til eru tvímælalaust kvenkyns, sbr. kvk.ft.**mgr**.

*eigendur**nar***, *fætur**nar***, *veturnar*, í stað upprunalegu karlkynsmyndanna kk.nf.ft.**mgr**. *eigendur**nir***, *fætur**nir***, *veturnir*, kk.þf.ft.**mgr**. *eigendur**na***, *fætur**na***, *veturna*. Auk þess eru stundum notaðar kvenkynsmyndir af ákvæðisorðum með karlkynsmyndum með ft. *-ur*. Í ljósi þess að nafnorð með endingunni ft. *-ur* geta verið hljóðfræðilega mjög ólík er í fyrstu greininni velt vöngum yfir því hvers vegna umtúlkunartíðni karlkynsmynda á ft. *-ur* sem kvenkyns komi ekki heim og saman við hlutfallslega tíðni karlkynsmyndanna sjálfra. Um þetta er staðhæft að búast mætti við samsvörun að öllu jöfnu. Samkvæmt leitarniðurstöðum úr textasöfnum fyrir íslensku er þó tíðni karlkynsmynda annars vegar og samsvarandi kvenkynsmynda hins vegar hlutfallslega ólík.

Í fyrstu greininni er beitt tengslanetsnálgun (e. *Network Model*) sem byggð er að mestu leyti á umfjöllun Bybee um verðandi (e. *emergent*) eðli málkefisins. Tengslanetin sem höfð eru í þessari ritgerð lýsa því hvernig hljóðfræðileg gerð kvenkynsundirflokksins í ft. *-ur* einkennist af samfellu (e. *gradation*). Þannig stuðla likindi við nafnorðin sem skipa X*ó*/*æ*T-örflokkinn að aukinni umtúlkunartíðni vegna áhrifa svokallaðs orðagengis (e. *lexical gang*). Áhrif örflokksins (e. *microclass*) eru álitin stafa af útvíkkun hliðstæðrar þekkingar, þ.e. ferlinu sem útvíkkar þekkingu til nýs samhengis.

Í svipuðum anda athugar önnur greinin takmarkaða virkni hins íslenska X*ó*/*æ*T-örflokksins en þar er sett samasemmerki milli virkni og beygingar tökuorðanna kvk. *blók* og hk. *kók* samkvæmt beygingarmynstrinu sem t.d. *bók* fylgir, sbr. et. *bók* ~ ft. *bæk**ur***, *blók* ~ *blæk**ur***, *kók* ~ *kæk**ur***. Talið er að hinar nýju fleirtölumyndir beri vitni um (að vísu mjög takmarkaða) virkni örflokksins, enda koma einnig fyrir fleirtölumyndir eftir beygingarflokkum kvenkynsorða sem hafa hærri mynsturstíðni, sbr. ft. *blók**ir***, *kók**ir***, eins og ft. *mynd**ir***. Fleirtölumyndirnar með endingunni *-ir* eru algengari en *blækur* og *kækur* en þetta er álitið sýna að tíðni og virkni haldist í hendur og að sambandið þar á milli ráði miklu um stefnu málbreytinga og fótfestuna sem nýjungar ná í málinu.

Önnur greinin bendir á að sumir álíta ólíkar hugrænar forsendur vera fyrir myndun fleirtölumyndanna ft. *blækur* og *kækur*. Í þessu sambandi hefur verið látið í veðri vaka að myndun ft. *kækur* endurspegli ekki „raunverulega“ málnotkun af því að viðkomandi mynd er aðeins notuð í gríni. Hins vegar sé ft. *blækur* dæmi um „raunverulega“ málnotkun. Gegn þessu viðhorfi sýnir önnur greinin fram á að báðar fleirtölumyndir hafa sama beygingarmynstur að fyrirmynd, þ.e. et. *bók* ~ ft. *bækur*. Af þessari ástæðu leitast önnur greinin við að hrekja þá hugmynd að ólíkar forsendur séu fyrir myndun fleirtölumyndanna sem í hlut eiga og að hugmyndin sjálf stafi af misskilningi á því hvernig málfræðileg þekking er útvíkkuð til nýs samhengis. Þessu markmiði er náð með vísun til hljóðfræðilegra gerða

ólíkra nafnorða, merkingar, málnotkunarfræði, auk niðurstaðna úr textasöfnum fyrir íslensku. Þannig eru nafnorð tekin inn í X*ó*/*æ*T-örflokkinn með samfelldum hætti (e. *in graded fashion*) og það er talið endurspegla takmarkaða virkni flokksins. Komist er að þeirri niðurstöðu að bæði ft. *blækur* og *kækur* séu myndaðar við útvíkkun hliðstæðrar þekkingar.

Þriðja greinin einblínir að útjöfnun (e. *levelling*) í beygingardæmum færeyskra karlkynsnafnorða, sem á fornvesturnorrænum tíma tilheyrðu flokki karlkyns *u*-stofna. Hér er útjöfnun talin hafa átt sér stað þegar stofnmynd tiltekins orð kemst inn í annan bás beygingardæmisins þar sem hana var ekki að finna áður. Lykilatriði í þessari skilgreiningu á utjöfnun er að ferlið felur ekki í sér breytingu eldri myndar í nýja, heldur viðbót við þær beygingarmyndir sem fyrir voru. Færeyskir afkomendur fornvesturnorrænna *u*-stofna hafa sætt mikilli útjöfnun en eins og beyging fær. *vøllur* ‘völlur’ sýnir er stofnmyndin *vøll-* (< fvn. nf.et., þf.et., þf.ft., þgf.ft. *vǫll-*) komin inn í allar beygingarmyndir orðsins þó að sumir básar hafi einnig stofnmyndina *vall-* (< fvn. ef.et., ef.ft. *vall-*) sem ekki höfðu hana áður. Aftur á móti er stofnmyndin fvn. þgf.et., nf.ft. *vell-* með öllu horfið úr beygingardæminu.

Aftur á móti hefur beygingardæmi nafnorðsins fær. *fjørður* ‘fjörður’ haldið öllum eldri stofnmyndunum sínum, sbr. fær. *fjørð-* (< fvn. nf.et., þf.et., þf.ft., þgf.ft. *fjǫrð-*), *firð-* (< fvn. þgf.ft., nf.ft. *firð-*), *fjarð-* (< fvn. ef.et., ef.ft. *fjarð-*). Enn fremur hafa allar stofnmyndir orðsins verið útvíkkaðar til básа innan beygingardæmisins þar sem þær komu ekki fyrir áður, sbr. nýjungarnar þgf.et. *fjørði*, yngri mynd sem lifir við hlið eldra *firði* og yngri myndirnar nf./þf.ft. *fjørðir* og *fjarðir* sem stundum leysa eldri *firðir* af hólmi og yngri myndirnar *fjarðum* og *firðum* sem eru í notkun ásamt eldri *fjørðum*.

Aðallega hefur verið miðað við tvær ólíkar kenningar um ákvörðun grunnmynda (e. *basic forms*), þ.e. þeirra mynda sem liggja formlega til grundvallar við myndun nýrra beygingarmynda innan beygingardæmis, og áhrifþættina sem stuðla að vali á grunnmyndum. Fyrri kenningin eignar valið stöðu mynda m.t.t. svokallaðrar mörkunar; samvkæmt því eru nýjar myndir sem til verða við útjöfnun taldar líklegri til að tjá merkingarfræðilega „eðlileg“ og/eða „hlutlaus“ gildi eins og *eintölu* og *nefnifall*. Á hinn bóginn gerir önnur kenning ráð fyrir að tíðni stuðli að vali á grunnmyndum, þannig að útjöfnun stefni út frá þeim myndum sem koma oftast fyrir í töluðu máli og/eða textasöfnum. Samkvæmt seinni kenningunni ákvarðar tíðni beygingarmyndar rótfestu (e. *entrenchment*) hennar, þ.e. hversu vel hugrænn fulltrúi viðkomandi myndar hefur búið um sig í minni. Með öðrum orðum eru algengar myndir rótfastari í minni en sjaldgæfari myndir og eru þær fyrrnefndu þ.a.l. tiltækari til notkunar en þær síðarnefndu þegar minnið bregst málnotandanum (þó ekki nema

stundarkorn). Þess vegna séu algengustu beygingarmyndir orðs líklegastar til að liggja til grundvallar við myndun nýrra mynda við útjöfnun.

Þriðja greinin styðst við innihald textasafna og beitir málnotkunarnálgun til að sýna fram á að grunnmyndir færeysku nafnorðanna *vøllur* og *fjørður* — sem bæði vísa til landslags og koma fyrir í samsettum staðanöfnum — má ákvarða á grundvelli tíðni, frekar en merkingar. Þar eru færð rök fyrir því að vegna staktíðni (e. *token frequency*) ólíkra mynda af orðinu *vøllur*, sem almennt er lág, hafi algengasta stofnmyndin, þ.e. *vøll-*, verið nýtt til grundvallar við myndun á nýjum beygingarmyndum en stofnmyndin *vell-* hafi síðan gleymst með tímanum. Aftur á móti varð há staktíðni myndarinnar þgf.et. *firði*, af nafnorðinu fær. *fjørður*, til þess að hún bjó vel um sig í minni og reyndist þ.a.l. tiltækari öðrum myndum þegar minnið brást. Af þessari ástæðu gerir þriðja greinin ráð fyrir að há staktíðni þgf.et. *firði* hafi jafnvel snemma stuðlað að útbreiðslu viðkomandi stofnmyndar til báss þágufallsmyndarinnar í fleirtölu í gegnum skemað (e. *schema*) [*í*/*á*/*úr* + þgf.].

Eftir að yngri myndin þgf.ft. *firðum* er orðin til tekur notkun hennar fram úr staktíðni eldri þgf.ft. *fjørðum*. Stofnmyndin *firð-* í fleirtöluhluta beygingardæmisins er komin inn í alla fleirtölubása, sbr. nf./þf.ft. *firðir*, þgf.ft. *firðum*, og tilvist hennar stuðlar enn fremur að sambandi forms og merkingar þar sem stofnmyndin *firð-* er gædd merkingunni *fleirtölu*. Gert er ráð fyrir að málnotendur bregðist síðan við þessu sambandi með tilraun til að jafna eintöluhluta beygingardæmisins út þannig að stofnmyndinni *fjørð-* sé skotið inn í bás þágufallsmyndarinnar í eintölu, þar sem nf.et. *fjørður* og þf.ft. *fjørð* var þegar að finna. Með öðrum orðum sé ætlunin að koma á andstæðunni et. *fjørð* : ft. *firð-*. En þrátt fyrir þessa viðleitni reynist eldri þgf.et. *firði* áfram algengari en yngri *fjørði* vegna staktíðni fyrrnefndu myndarinnar og þótt sú síðarnefnda lifi enn í málinu er hún enn tiltölulega sjaldgæf.

Ritgerð þessi sýnir fram á að þættir á borð við tíðni og líkindi við hljóðfræðilega gerð annarra orðmynda geta mörgu ráðið hvað varðar málnotkun, auk þess að þessir áhrifaþættir eiga sér rætur í fyrri reynslu við málið. Enn fremur eru færð rök fyrir því að haldist stefna málbreytinga í hendur við áðurnefnda þætti þá sé rökrétt að líta svo á að málkerfið sé í stöðugri mótun á grundvelli sömu þáttanna. Með öðrum orðum renna rannsóknirnar sem hér eru til umfjöllunar stoðum undir það sjónarmið að málkerfið sé verðandi (e. *emergent*) fyrirbæri frekar en að mótun þess sé takmörkuð á grundvelli líf- og erfðafræðilegra þátta. Þar af leiðandi og í ljósi niðurstaðnanna sem greinarnar þrjár varpa fram má líta á þessa ritgerð sem innlegg í innlenda og alþjóðlega umræðu um áhrif almennra hugrænna þátta á málnotkun og stefnu málbreytinga, auk þess að sýna nytsemi málnotkunarnálgunar í rannsóknum á beygingarbreytingum í eyjanorrænu.

Table of contents

Acknowledgments

Completion of the current thesis would not have been possible without the guidance of Þórhallur Eyþórsson (my supervisor), Hjalmar P. Petersen, and Katrín Axelsdóttir (doctoral committee members). I owe them the greatest debt of thanks for assistance with everything from funding to submission of the final product. I also express profound thanks to former teachers: Daisy L. Neijmann, Guðrún Þórhallsdóttir, Gunnþórunn Guðmundsdóttir, Jón Axel Harðarson, Kristján Árnason, and Margrét Jónsdóttir. I wish to acknowledge the input of my friends and colleagues in the Department of Icelandic as a Second Language (University of Iceland), who have provided support and encouragement, in addition to showing understanding when the call of the thesis required my withdrawal into the shadows.

The route to completion would not have been as scenic without the impact of my wonderful friends from Aarhus University: Peter Bakker, Kristoffer Friis Bøegh, Joshua Nash, and Jeroen Willemsen. Thank you for making me feel like I belonged to the department in 2018 and for your continuing friendship. For the same reasons, I am grateful to Turið Sigurðardóttir and all at Føroyamálsdeild, University of the Faroe Islands, for accommodating me on my regular work visits from 2017–2019. I also owe a great debt of thanks to my friend Charles Gittins, for his invaluable assistance with formatting before submission.

On a personal note, I wish to express gratitude and love to my Danish family: Margit and Søren, Thomas and Annabel, Rosalina, and Solveig. And to my Icelandic family: Anna, Björg, Oddrún, Óli, Ólöf (Sandra Hermannsdóttir); Ólöf (Matthíasdóttir) and Skúli; Brynja (amma); Erna Björk. I thank my dear friend, Gabríela, for showing in 2015 that all creative work is art. Thank you for accepting me into your lives and for becoming part of mine.

I shall ever be lucky to have known nothing but unfaltering love, acceptance, and support from my family: Angela Irving, Mark Cretton, Paul Mariner, Christian Cretton, Rebecca Mariner, John Irving, Beckie Cretton, and bellends 1–4. To my beautiful Jesper. This process would literally not have been possible without you and the bubbas by my side. I dedicate this thesis to the memory of Anne Irving and John & Pat Cretton, my dearly departed grandparents.

With love,
Jón Símon

Work on this thesis was funded by The Icelandic Research Fund, grant no. 174253-015.

1 Introduction

This chapter introduces the focus of the thesis, setting out the material chosen for analysis within the context of the theoretical approach employed. The contents of the current chapter are as follows. In Section 1.1, I flesh out the thematic focus and state my objectives. In 1.2, I reiterate the individual research questions and specific aims as stated in the respective studies, which form the basis of the thesis. Section 1.3 presents the general contents of the thesis, with respect to the order of subsequent chapters.

1.1 Focus and general research questions

The current thesis elaborates on three original articles, each of which focuses on morphological change in Icelandic or Faroese. Referred to collectively in the current thesis as Insular Nordic, the two languages constitute West Nordic sisters within the North Germanic branch of the Germanic language family (see e.g. Vikør 2001:32–33). Throughout, the articles are referred to as Paper I (PI), Paper II (PII), and Paper III (PIII), respectively.[1] Each article has been published in a well-established, peer-reviewed journal of linguistics. The precise bibliographical entries for Papers I–III are given below.

I. Markússon, Jón Símon. (2023a). Accounting for different rates of gender reanalysis among Icelandic masculine forms in plural *-ur*. *Nordic Journal of Linguistics*. <https://doi.org/10.1017/S0332586522000166> [Published online 2022.]

II. Markússon, Jón Símon. (2022a). Tvær blækur labba inn á bar: On limited productivity as graded membership of an Icelandic microclass. *NOWELE*, 75(2), 194–222. <https://doi.org/10.1075/nowele.00068.mar>

[1] When referring to a specific **page** in an article, the notation e.g. **PI:5** will be used and is to be read 'On page 5 of Paper I'. When referring to an **endnote** or **footnote** within an article, the notation e.g. **PII:201, footnote 6** will be used, to be read 'footnote 6 in Paper II, which occurs on page 201 of the same article'. The reader is reminded that Paper I contains endnotes, rather than footnotes. When referencing a **table** in an article, the notation e.g. **PIII:65, Table 3** will be used and should be read 'Table 3 in Paper III, which occurs on page 65 of the same article'.

III. Markússon, Jón Símon. (2022b). Um áhrif tíðni á stefnu útjöfnunar: Rannsókn á beygingarþróun færeysku nafnorðanna *vøllur* og *fjørður*. *Íslenskt mál og almenn málfræði*, 44, 53–86.

As can be discerned from the bibliographical information above, Papers I and II are written in English, while Paper III is in Icelandic. Each study was conducted in accordance with the principles and axioms of the usage-based cognitive approach to language change.

Papers I and II discuss different inflection classes of Icelandic and the limited productivity associated with some of these. In particular, both articles account for productivity associated with the Icelandic ending nominative/accusative plural *-ur* (henceforth 'plural *-ur*') and feminine subclasses that show that ending. Specifically, Paper I accounts for the different rates at which masculine forms in plural *-ur* are reanalysed as feminine due to its high rate of **dispersion** among Icelandic feminine nouns. Throughout, *dispersion* refers to the distinct morphological contexts in which a given exponent occurs (following Gries and Ellis 2015:232). Paper II deals with graded membership of the borrowings Ice. fem. *blók* 'non-entity, wretch' and. neut./fem. *kók* 'Coke™' in several feminine classes as a function of **minimal schematicity**, i.e. a high degree of phonetic similarity to the members of a given class.[2]

In line with e.g. Bybee (2010), the gauge for **productivity** is the number of items to which an inflectional pattern is extended. Further, in line with Barðdal (2008:1), it is argued that productivity is in part determined by the number of items *attracted* to a given inflection class on the basis of phonetic coherence with existing members. In other words, the measure of productivity is also predicated on schematicity, a property of an inflection class that correlates with the **extensibility** of the relevant inflectional pattern. It has been demonstrated that those inflection classes of a language that contain the most members and whose membership is characterised by high schematicity prove the most productive. This property is considered a function of the ease with which highly frequent patterns are relatively more practiced and, therefore, more easily applied in new contexts. Likewise, a lack of formal constraints on a class of items facilitates extension of membership to new items of practically any form (e.g. Bybee 2010:67).

However, morphological productivity is a demonstrably graded phenomenon and Papers I and II are concerned with the limited productivity that can be exhibited by classes

[2] Throughout, I follow Bybee's (2010:67) definition of **schematicity**, i.e. the degree of formal *dissimilarity* between the members of a class (see 3.2.3 for an in-depth discussion).

and subclasses that contain fewer members. In accordance with the graded view, limited productivity is considered a function of low to minimal schematicity, i.e. when the members of a given class fit a tight phonetic definition. It is argued that this property of a class can offset the limiting effect of diminutive class size (see above). Specifically, Papers I and II demonstrate that small classes of Icelandic nouns are able to attract new members by means of a **gang effect**, defined in terms of the attractive force facilitated by a high concentration of common phonetic attributes among the members of a small inflection class (Bybee 2010:69; see 3.2.3; also Axelsdóttir 2015 and Markússon 2021, 2022a, 2023a on the gang effect in Icelandic and Faroese).

Paper III accounts for the inflectional development of the Faroese nouns *vøllur* and *fjørður*. The former has in the main levelled out alternation between old stem variants, cf. Old West Nordic (OWN) nom./acc.sg., acc./dat.pl. *vǫll-*, dat.sg., nom.pl. *vell-*, gen.sg./pl. *vall-*. As a result of **analogical levelling**, defined as the process by which a paradigm acquires new forms based on existing stem variants (Carstairs-McCarthy 2017:327), all paradigm cells for Far. *vøllur* now contain the variant *vøll-* (< OWN *vǫll-*), while the variant *vall-* has also spread to a much lesser extent. Consequently, in the main, *vøllur* has joined the ranks of a larger class of masculine nouns; one in which stem alternation was rare and whose nominative/accusative plural ends in *-ir* in Modern Faroese, cf. Far. nom.sg. *vøllur* ~ nom./acc.pl. *vøll**ir*** (< OWN nom.pl. *v**e**llir*), *gestur* ‘guest’ ~ *gest**ir***.

Conversely, Far. *fjørður* has not only retained the reflexes of all its old stem variants, cf. OWN nom./acc.sg., acc./dat.pl. *fjǫrð-* (> Far. *fjørð-*), dat.sg., nom.pl. *firð-*, gen. sg./pl. *fjarð-*, but also has extended them to other cells of the paradigm. Significantly, the paradigm represents the only Old West Nordic *u*-stem to exhibit such development in Faroese language history. In light of the different paths of levelling discussed above, Paper III argues that the historical token frequencies of individual forms, in addition to a stem variant’s historical dispersion, determined the direction of change. Further, the different frequency effects elucidated in Paper III give cause to question the approach to levelling associated with so-called ‘markedness’.

Individually, Papers I–III account for various types of analogical change. As noted above, the changes discussed here are considered a function of productivity. In this connection, analogy is posited as the cognitive means via which productivity proceeds. For example, Paper II posits **analogical extension** as the mechanism of change, the motivation for which is perceived phonetic similarity to functionally equivalent relations in other paradigms (Bybee 2015:93–94). Thus, it is argued that phonetic similarity between borrowed

Ice. *blók* and a microclass that originally contained only *bók* ‘book’, *bót* ‘patch, remedy’, *brók* ‘trousers’, *nót* ‘(fishing) net’, and *rót* ‘root’ — referred to here as the **X*ó*/*æ*T-microclass** — facilitated categorisation of the borrowing with the latter set, cf. *bók* ~ plural *bækur*, *bót* ~ *bætur*, *brók* ~ *brækur*, *nót* ~ *nætur*, *rót* ~ *rætur*, thus (borrowed) *blók* ~ (innovative) *blækur*.[3]

Conversely, Paper II argues that, while the innovative alternation Ice. sg. *kók* ~ plural *kækur* stems from extension of the pattern exhibited by e.g. *bók* ~ *bækur*, occurrence of fem. dat.sg. *kók*, as opposed to neut. dat.sg. *kóki*, stems from levelling (see PII:206). Although extension and levelling are viewed as variant mechanisms of proportional analogy, levelling is characterised by an attempt at eliminating “unnecessary” alternation between forms of the same word. In other words, through levelling, stem variants from within an inflectional paradigm typically replace older stem variants. Replacement can occur on occasion or, eventually, one variant can be completely replaced. Crucially, the form of new variants is based on existing forms *within* the relevant paradigm.

As demonstrated by Bybee (1985), the structure of inflectional paradigms across multiple, unrelated languages is often characterised by phonetic similarity among those forms of a word which express a common morphological value of significant semantic weight. This property of paradigms is reflected in the direction of levelling. Among nouns, the values *singular* and *plural* (and *dual*, if the relevant language makes that distinction) hold significant semantic weight relative to each other: A single entity is not the same thing as multiple of the same entity. Thus, it is argued that syncretic fem. nom./acc./dat.sg. *kók* reflects an attempt to align meaning and form, as syncretism better reflects common semantic content than does distinction between nom./acc.sg. *kók* and dat.sg. *kóki*. As the variant dat.sg. *kók* is based on a form found elsewhere in the paradigm, deduction of the form is considered the result of levelling.

However, despite substantive differences between extension and levelling (see above), syncretism in nom./acc./dat.sg. *kók* accords with a pattern that is highly characteristic of the inflection of strong feminine nouns,[4] cf. fem. nom./acc./dat.sg. *bók*, *mynd*. Therefore, it is

[3] The notation X*ó*/*æ*T-microclass is actually used in Paper I, while the same set of feminine nouns is referred to as the X*ó*T-microclass in Paper II. I have decided to use the former notation in the extended introduction, as it is more informative as to the pattern of alternation between sg. *-ó-* and nom./acc.pl. *-æ-*, according to which members of the microclass conform, cf. e.g. the relations *b**ó**k* ~ *b**æ**kur*, *r**ó**t* ~ *r**æ**tur*.

[4] Strong Icelandic nouns are those whose genitive singular ends in a consonant; weak nouns are those whose genitive singular ends in a vowel. This applies to nouns of all three genders (Kvaran 2005:221).

argued that levelling to dat.sg. *kók* and the form's substitution for dat.sg. *kóki* demonstrates external pressures on relations within the paradigm of Ice. *kók*. In other words, the model of an existing pattern motivates levelling via extension, by which neut. sg. nom./acc. *kók* ~ dat. *kóki* is levelled to fem. nom./acc./dat. *kók*. The distinction between the two mechanisms of proportional analogy is, therefore, not completely clear cut as substantive crossover is often clearly observable.

Similarly, Paper III argues that masculine classes of higher **type frequency**, defined as the number of individual words — each counted once — that conform to a pattern, exerted pressure on the paradigms of Old West Nordic *u*-stems (see above). However, the survival of stem-variants in the development of Far. *fjørður* is accounted for by reference to other measurements of frequency also. One of these measurements is conducted along the dimension of **token frequency**, defined as the number of occurrences of an inflectional form — each counted once — in spoken language or running text.

Thus, for example, as Far. *fjørður* is a highly frequent latter constituent of complex place names that designate both a fjord and human settlement, e.g. *Fuglafjørður*, *Kollafjørður*, it occurs frequently in the dative case, which in Faroese performs the same function as the locative case in e.g. Russian, as well as others (see Þráinsson et al. 2012:164 on the dative case on objects of the prepositions *á* 'on', *í* 'in', and others).[5] Conversely, due to attributes of the Faroese landscape, the topographical referents of Far. *vøllur* tend to be very small and, therefore, not to perform established functions. Indeed, the study reported in Paper III reveals that the dative form of Far. *fjørður* is of high token frequency, while occurrence of Far. *vøllur* is characterised by low token frequency in all of its case forms. In other words, language users' perception of real-world conditions appears to determine the frequency of words and their individual inflection forms (following Haspelmath 2006).

From the usage-based cognitive perspective, high frequency is a correlate of **entrenchment**, defined as the reorganisation of knowledge in memory on account of the impact of continued experience (Schmid 2017:3–4). Thus, the more frequently an individual experience a phenomenon in the world, the more embedded, i.e. entrenched, the cognitive representation of that experience becomes relative to the representations of less frequent phenomena. In turn, entrenchment is viewed as the determinant of **lexical strength**, i.e. how accessible information pertaining to the relevant experience is from memory (Bybee

[5] See also Stolyarova 2016, who discusses and compares the equivalent roles of the locative and dative cases in Russian and Icelandic, respectively.

1985:117). In this spirit, Paper III argues that the original stem of dat.sg. ***firð****i* is more entrenched than that of original dat.pl. ***fjørð****um* in syntactic contexts such [*í*/*úr* FJØRÐUR$_{dat.}$], as place names with *-fjørður* almost always contain the singular form.[6] This property of usage is a function of token frequency, which facilitated greater lexical strength of the stem variant (dat.sg.) *firð-*, a position supported by the fact that younger dat.pl. *firðum*, based on the high-frequency dative singular form, is now much more frequent than older dat.pl. *fjørðum*.

Paper III subsequently argues that the deduction of innovative dat.sg. *fjørði* represents an attempt to contrive a dynamic whereby the stem variant *firð-* is now synonymous with *plural*, cf. nom./acc.pl. ***firð****ir*, innovative dat.pl. ***firð****um*, opposing singular forms with *fjørð-*, cf. nom.sg. ***fjørð****ur*, acc.sg. ***fjørð***, dat.sg. ***fjørð****i*. It is argued that this formal opposition is motivated by the dispersion of stem variants throughout the paradigm (see above). However, dat.sg. *firði* has most probably always had far higher token frequency than all other forms of *fjørður* and its stem variant will have been more accessible than others as a result. This property of use renders the likelihood that the variant *firð-* be forgotten — albeit only momentarily — very unlikely, a position that is supported by the fact that use of younger dat.sg. *fjørði* is characterised by low token frequency, meaning it has never acquired the degree of entrenchment likely required to usurp its older cellmate's place in memory.

In light of the brief overview above, the general objectives of the thesis are set out in (1).

(1) a. To determine which cognitive factors have impacted the structure of the inflectional systems of Icelandic and Faroese.

b. To ascertain whether these cognitive factors apply only to language or are domain-general in nature.

c. To demonstrate how the impact of these cognitive factors are evident through language change as a function of language use.

[6] The prepositions Far. *í* 'in' and *úr* 'from' both assign dative to their objects to convey location. The appropriate forms of Far. *fjørður* within this context would have been dat.sg. *firði* and dat.pl. *fjørðum* before the process of levelling began (see PIII:76–80).

In order to fulfil the objectives stated in (1), subsequent chapters attempt to provide answers to the respective research questions posed in Papers I–III, which are reiterated in the next section.

1.2 Aims and individual research questions of Papers I–III

This section sets out the stated aims of Papers I–III and lists the research questions of each study in the context of those aims. The aim of Paper I is first and foremost to examine and account for the extent to which schematicity impacts the rate at which Icelandic masculine forms in plural *-ur* are reanalysed as feminine. The research statement of Paper I (PI:1) is restated as the question in (2).

(2) How do we account for the different rates at which Icelandic masculine forms in plural *-ur* are treated as feminine due to varying degrees of phonetic and/or semantic similarity to clusters of feminines in the same ending?

Paper I argues that due to the high rate at which the ending plural *-ur* is dispersed among feminine nouns, any masculine form in the same ending can be reanalysed as feminine — all other things being equal. However, it is also demonstrated that phonetic coherence with the X*ó*/*æ*T-microclass of feminines in plural *-ur* boosts the rate of reanalysis.

Paper II aims specifically to account for the limited productivity of the Icelandic X*ó*/*æ*T-microclass, discussed in Section 1.1. The objective of Paper II (see PI:196) is stated in (3).

(3) To provide a usage-based cognitive account of the limited productivity of the Icelandic X*ó*/*æ*T-microclass through recourse to:

i. Semantics,
ii. Pragmatics,
iii. Schematicity,
iv. Analogy.

In this connection, the productivity of the microclass is betrayed by the addition of two loans from English, namely *blók*, cf. British/Australian Eng. *bloke*, and *kók*, cf. Eng. *Coke*™. Paper II argues that assignment of the loans to the X*ó*/*æ*T-microclass is a function of the gang

effect. However, the analysis demonstrates that a property of limited productivity is graded membership of the microclass, whereby inflection of the loans only sometimes accords with the relative pattern, and sometimes with patterns that characterise other classes. Paper II argues that the respective degrees to which each noun is incorporated in the X*ó*/*æ*T-microclass are determined by their individual semantic properties, as well as use in context, i.e. pragmatics.

The aim of Paper III is to identify the distinct inflectional forms within the paradigms for Far. *vøllur* and *fjørður* from which levelling proceeds. As part of this aim, the analysis also endeavours to ascertain the role of frequency and/or semantics as determinants of the direction of levelling. The research question of Paper III (PIII:54) is reiterated in (4).[7]

(4) Does frequency determine the direction of levelling?

Paper III argues that the direction of levelling is determined by token frequency as a correlate of entrenchment and lexical strength.

1.3 General contents of the thesis

Chapter 2 discusses the genetic relationship between the West Nordic languages, Icelandic, Faroese, and — to a lesser extent — Norwegian, as well as inflectional change in all three languages since the Old West Nordic period. Chapter 3 fleshes out the usage-based cognitive approach to language change. There, I justify application of the approach to the Icelandic and Faroese data presented in Papers I–III on the basis of comparison with the approach from generative grammar and dual-processing. Chapter 4 delineates the methodology employed in Papers I–III in light of the theoretical tenets of the usage-based cognitive approach, as outlined in the previous chapter. Chapter 5 discusses the specifics of Papers I and II in terms of thematic content, application of the usage-based cognitive approach to the relevant Icelandic data, and the conclusions drawn in each paper. Chapter 6 provides a detailed usage-based cognitive analysis of the Faroese data presented in Paper III and relevant conclusions. Chapter 7 draws general conclusions through reference to the research questions set out in (1) (Section 1). Finally, access is provided to Papers I, II, and III at the end of the extended introduction, under the heading ‘Paper I’, ‘Paper 2’, and ‘Paper III’.

[7] Translated from Ice. *Er tíðni ráðandi afl í stefnu útjöfnunar?* (PIII:54).

2 Insular Nordic

In this Chapter, I discuss the genetic and historical relationship between the West Nordic languages with particular focus on Insular Nordic and, to a lesser extent, through reference to Norwegian. The discussion is by no means exhaustive but is intended as an overview. Below, I draw comparisons between Old West Nordic and Insular Nordic with the specific intention of demonstrating the different degrees to which the constituent inflectional systems of the latter have diverged from the former. Further, I elucidate the nature of the changes discussed. The content of this section will feed into Chapters 3 and 4, where I discuss two opposing theoretical approaches to language change and the methodology employed in Papers I–III, respectively.

Along with the dialects of Western Norway, as well as the now extinct Orkney and Shetland Norn, Icelandic and Faroese belong to the West Nordic branch of the Nordic language family tree.[8] Collectively, Icelandic and Faroese furnish the Insular Nordic branch, due to their distribution across geographically distinct Nordic island nations, i.e. Iceland and the Faroe Islands. The Insular Nordic languages both derive from Old Norse through Old West Nordic. From a period spanning roughly the 9th to the 11th centuries, loosely designated as the "Viking Age", the seafaring, Norse-speaking peoples of Mainland Scandinavia — particularly Western Norway — sought out new lands to the South-West and West. The fruits of this exploration were subsequent settlements on Orkney, Shetland, in Scotland and the Hebrides, in the Faroe Islands, Iceland, Greenland, and in North America. In each settlement, a variety of Nordic was the community language (e.g. Vikør 21–23). While initial Nordic presence in America was short-lived, a Nordic community in Greenland survived in co-existence with Inuit until the former seemingly disappears before the early 16th century. Conversely, the West Nordic dialects of Orkney and Shetland survived until the mid-1700s and just prior to 1800, respectively (Barnes 2004:75). Since Norn went extinct, Icelandic and Faroese have stood alone as the only extant descendants of Old West Nordic to the West of Mainland Scandinavia.

In the context of the developments just delineated, questions arise as to the extent to which Icelandic and Faroese have each diverged from Old West Nordic and to which each

[8] Some dialects of Norwegian, e.g. that spoken in and around Oslo, as well as Danish and Swedish (formerly also Old Gutnish), occupy the East-Nordic branch (e.g. Vikør 2001:33).

language resembles the another. In this connection, change in Insular Nordic is documented to starkly different degrees in the written record. The earliest attestations of Icelandic date from roughly 1150–1200, from which time the language is consistently well attested until the present day. Due to the sheer volume of written sources for Icelandic, historical linguists have been able to plot the chronology of change to a significant degree of precision relative to diachrony.

Conversely, Faroese language history is sparsely documented until around 1800 (Þráinsson et al. 2012, e.g. p. 374). Early written sources for the language include two runic inscriptions found in the Faroe Islands: one in the village of Kirkjubøur and the other in Sandavágur, dated to roughly 1000 and 1200, respectively. Probably the most important early Faroese source in terms of comprehensiveness constitutes Far. *Seyðabrævið* 'The Sheep Document', dated to 1298. The document bestows regulations for keeping sheep in the Faroe Islands and the original is believed to contain linguistic features that developed in the Faroes. Further, a transcript of *Seyðabrævið* from around 1600 is considered to reflect hints of a budding new variety of West Nordic, i.e. Early Faroese (see e.g. Þráinsson et al. 2012:371).

It is written documents from the beginning of the 15th century which first show that a clearly distinct Faroese variety of West Nordic has developed. The main sources for this variety are the Far. *Húsavíkarbrøvini* 'The Húsavík Letters', which exist in transcribed form from 1407 and 1479. Here, spellings such as *hrentadi* for older *rentaði* suggest that scribes were unsure where to write *hr* for older initial voiceless *r*, a West Nordic trait that had likely disappeared from Faroese by the 15th century or was almost lost at the time of transcription (see Þráinsson et al. 2012:372). Such spellings suggest a clear difference between Faroese and Icelandic: voiceless *r* still occurs word-initially in the latter and is represented by orthographic *hr*, e.g. Ice. ***hr****ingur* 'ring', ***hr****einn* 'clean, pure', ***hr****ópa* 'yell, call'.

Due to developments such as those just discussed, historical linguists commonly characterise change in Faroese as more progressive than in Icelandic. In other words, the latter is considered more conservative than the former. A common observation in this regard is that the inflectional system of Modern Icelandic better resembles that of Old West Nordic than does Modern Faroese. In turn, the latter's inflectional system harkens more so to that of Old West Nordic than does Modern Norwegian and the other Mainland Scandinavian languages, i.e. Bokmål, Danish, and Swedish (e.g. Þráinsson et al. 2012:369). Despite this, it is reasonable to assert that many changes observed in Icelandic have also occurred in Faroese, while change in the latter has led to highly complex intra-paradigmatic dynamics not mirrored in the inflectional system of Icelandic (see below).

While there is a marked difference in the degree of conservatism that Icelandic and Faroese exhibit to Old West Nordic, there are also conspicuous structural similarities between both Insular Nordic languages that distinguish them from Mainland Scandinavian. This is particularly true in the context of morphology but also in (morpho)syntax (e.g. Þráinsson et al. 2012:59, 407, Eyþórsson 2015). For example, while the respective phonological systems of Insular Nordic have diverged from that of Old West Nordic and each other, both Icelandic and Faroese retain a complex and active inflectional system for nouns. One commonly cited non-linguistic reason for this conservatism is the geographical isolation of both island nations through the centuries, while another is the kinds of social dynamics that typically characterise small, insular communities. In terms of degree, then, such factors might both account for the retention of morphological complexity in both languages but also for the fact that changes to the inflectional system of Faroese have led to more complexity than in that of Icelandic (see e.g. Trudgill 2011, who discusses how such socio-linguistic dynamics can affect language structure).

In other words, while the Faroese system of inflection has changed more than that of Icelandic (see below), intra-paradigmatic change has not been synonymous with simplification, thus counter to ideas presented by Kiparsky (1974).[9] For example, the morphologisation of historically phonological processes such as *verschärfung* and palatalisation in Faroese has served to complexify intra-paradigmatic dynamics as these pertain to alternation between inflectional forms (Petersen 1992; also Petersen and Voeltzel, to appear 2024). Thus, the paradigm of the Icelandic verb *búa* ‘live’ bears witness to *verschärfung* but only in the past plural, cf. pres. 1sg. *bý*, 2sg. *býrð*, 3sg. *býr*; past 1/2/3sg. *bjó-*; past 1/2/3pl. *bju**gg**-*. However, in the paradigm of the Faroese cognate, *búgva* ‘live’, we see the result of *verschärfung* in different person/number combinations in the present: pres. 1sg. *bú**gv**i*, 1/2/3pl. *bú**gv**a* but pres. 2sg. *býrt*, 3sg. *býr*; but not in the past, past sg. *búði*, past pl. *búðu*. Thus, unlike the Icelandic dynamics delineated above, *verschärfung* in the Faroese paradigm is not at all associated with the past and not fully with the present, nor with a specific person/number combination (see Weyhe 1996, pp. 81–82, for examples of complexity resulting from the morphologisation of palatalisation).

The more conservative nature of the Icelandic inflectional system is best characterised in relation to the degree of change exhibited by the respective systems of Faroese and Norwegian. Thus, for example, while Icelandic retains the four-case system for nominals it

[9] See Chapter 3 for further discussion on the notion of intra-paradigmatic change as simplification.

inherited from Old West Nordic — inflecting for nominative, accusative, dative, and genitive — Faroese has lost its genitive case.[10] As the result of more sweeping change, Norwegian nominals, bar remnants among the personal pronouns, are generally not inflected for case, although genitive *-s* has been retained as a marker of possession.[11] While Icelandic and Faroese retain three distinct grammatical genders for nominals — masculine, feminine, and neuter — the distinction between masculine and feminine is no longer maintained in all varieties of Norwegian, yielding a class of *common gender* nouns that opposes the neuter in a two-gender system (Vikør 2001:38, 207–209).

Despite the relative conservatism of Icelandic, the system by which its nominals inflect has by no means been left unaltered. To offer some random examples, the Old West Nordic *u*-stems eventually lost the ending acc.pl. *-u* to the *i*-stem ending acc.pl. *-i* in the early 16th century (Þórólfsson 1925:22), cf. OWN *u*-st. acc.pl. *fjǫrðu* → *firði* 'fjords'. Almost all members of the feminine class of *ijō*-stems have lost the older ending nom.sg. *-(u)r* due to its association with masculine grammatical gender, e.g. OWN fem. *ermr* → Ice. *ermi* 'sleeve', although cf. the women's names OWN/Ice. fem. *Auð*(***u***)***r***, *Sigríð*(***u***)***r***, *Þuríð*(***u***)***r***, also masc. OWN *hestr* > Ice. *hestur* 'horse' (e.g. Þórhallsdóttir 1997). Levelling has affected disyllabic masculine stems with the suffix *-il-*, cf. e.g. OWN nom.sg. *l**y**kill* 'key' ~ dat.sg. *l**u**kli* ~ nom.pl. *l**u**klar*, Ice. nom.sg. *l**y**kill* ~ dat.sg. *l**y**kli* ~ nom.sg. *l**y**klar*, although OWN/Ice. nom.sg. *k**e**till* 'kettle' ~ dat.sg. *k**a**tli* ~ nom.pl. *k**a**tlar*.

Many changes to the inflectional system of Icelandic are mirrored in Faroese. Thus, for example, the latter has also generalised the root vowel *y* throughout the paradigm of Far. *l**y**kil* 'key', cf. dat.sg. *l**y**kli*, nom./acc.pl. *l**y**klar*. However, typically, change of this nature has gone even further in Faroese, cf. Far. nom.sg. *k**e**til* 'kettle', dat.sg. *k**e**til*, nom./acc.pl. *k**e**tlar*; even eradicating alternation that resulted from Proto-Nordic *u*-umlaut,[12] which has been preserved

[10] The genitive is no longer actively assigned in Faroese. Rather, use of old genitive forms appears to be fully conventionalised in phrases such as Far. *til Føroya* 'to the Faroe Islands', where the preposition *til* assigned genitive in Old West Nordic (and still does in Icelandic). However, synchronically, objects of *til* in the modern language are accusatives. While genitive forms occur in new compounds, e.g. *bending**ar**grunnur* 'inflection database', where the element *-ar-* is the old genitive ending, these are modelled on older genitive compounds such as *bending**ar**frøði* 'inflectional morphology' and function as linking elements (see Jónsson and Eyþórsson 2011 for a comparison of the development of genitive arguments in Icelandic and Faroese, and Petersen and Szczepaniak 2018 on genitive forms in the broader Germanic context).

[11] Some dialects of Eastern Norway retain dat.pl. *-um*, while some Western dialects have *-o*/*-å*. This is exceptional, however (see Sjekkeland 1997, Eyþórsson et al. 2012).

[12] On the nature of *u*-umlaut in Icelandic and Faroese language history, see Markússon 2017 and sources cited there.

in Icelandic, e.g. OWN nom.pl. *katlar* ~ dat.pl. *kǫtlum*, Ice. *katlar* ~ *kötlum*, cf. Far. ***k*e*tlar** ~ ***k*e*tlum***. The ending of OWN nom.sg. ***erm*r*** has also been replaced in Faroese, although by *-a* rather than *-i*, cf. Far. fem. ***erm*a*** 'sleeve', Ice. fem. ***erm*i***. Again, typically, Faroese has taken change to Old West Nordic *ijō*-stems even further, cf. the Faroese women's names *Sigrið*, *Turið* and their Icelandic cognates, *Sigríður*, *Þuríður*.[13]

Sweeping processes of levelling have occurred in Faroese, which lack equivalents in Icelandic. For example, alternation stemming from Proto-Nordic *u*-umlaut has in the majority of cases been levelled in the Faroese reflexes of Old West Nordic *ōn*-stems — all of which were feminine nouns. Thus, alternation between OWN *a* and *ǫ* has mostly been levelled in favour of Far. *ø* (< OWN *ǫ* by regular sound change), cf. e.g. OWN nom.sg. *gata* 'path, road' ~ acc./dat./gen.sg. *gǫtu*, Ice. *gata* ~ *götu*, but Far. ***gø*ta** ~ ***gø*tu** 'street'; OWN *saga* 'history, saga, story'~ *sǫgu*, Ice. *saga* ~ *sögu*, but Far. *søga* ~ *søgu*. Conversely, some nouns that follow the affixal pattern of Faroese *ōn*-stems show the stem-vowel *a* in all forms, e.g. Far. nom.sg. ***ma*mma** 'mum, mummy' ~ acc./dat.sg. ***ma*mmu**, cf. Ice. *mamma* ~ *mömmu* (see Markússon 2017).

Despite the kind of simplification by levelling exemplified above, the process has also yielded complexity in Faroese paradigms. Here, it is necessary to note that, in essence, levelling does not change one inflectional form into another. Instead, the process is defined as the deduction of a new form based on the (phonetic or morphological) structure of another within the same paradigm. In other words, functionally equivalent doublets that existed before levelling are not suddenly eradicated from the inflectional paradigm when a new form is conceived. Rather, these may co-exist for centuries (e.g. Bybee 2015:95).

Examples of complication by levelling can be taken from the paradigms of Faroese nouns containing either *a* or *ø* in the stem of the nominative singular. Thus, the paradigms og masc. *hvalur* 'whale' and fem. *øksl* 'shoulder' contain the dative plural forms ***hvø*lum** and ***ø*kslum**, respectively, where *ø* is the regular reflex of Proto-Nordic (PNc.) **a* by the historical process of *u*-umlaut, cf. PNc. dat.pl. **hv*a*lum*(ʀ) > OWN *hvǫlum* > Far. ***hvø*lum**, PNc. dat.pl. ***a*xlum*(ʀ) > OWN ***ǫ*xlum** > Far. ***ø*kslum**. However, in the course of Faroese language history, both paradigms have acquired *ø*-less forms, cf. dat.pl. ***hva*lum** and ***a*kslum**, based on the stem

[13] According to Petersen (2009), change from nom.sg. *-r* to *-a* and/or the lack of an ending was the only available choice to speakers of Faroese, whose aim was retaining feminine grammatical gender in e.g. *erma* and *Sigrið*. This, Petersen argues, is because Far. nom.sg. *-ur* (< OWN *-r*) and *-i* are overtly masculine markers. Therefore, one could argue that, compared with the developments of Icelandic cognates, the formal parameters for gender distinction in Faroese are far clearer cut.

variants *hv**a**l-* and ***a**ksl-* from elsewhere within the respective paradigms, e.g. nom.sg. *hv**a**lur*, nom./acc.pl. ***a**kslar*/***a**kslir* (see Þráinsson et al. 2012:398–399). While the forms with *a* might constitute instances of simplification, their coexistence with forms with *ø* add complexity that formerly did not exist in the respective paradigms.[14]

While much analogical change in Faroese can be characterised as levelling, Icelandic paradigms have undergone the process to a far lesser extent. For example, the paradigms of Icelandic feminine nouns in nom.sg. *-a* have retained *u*-umlaut alternation that parallels the relational oppositions characteristic of *ōn*-stem inflection in Old West Nordic (see examples above). Likewise, while alternation between individual stem variants of Old West Nordic *u*-stems has been all but levelled out in Faroese (see 1.2), the inflection of Icelandic *u*-stems strongly reflects the pattern of alternation that characterised their Old West Nordic predecessors.

This brief overview of change in Insular Nordic nouns demonstrates that Faroese has diverged more from Old West Nordic than has Icelandic, and that the kinds of change witnessed between the two languages are often of a substantively different nature. Thus, change in Faroese provides ample opportunity for research into the nature of levelling. Conversely, Icelandic is more conservative with regard to the retention of ancient vocalic alternation, more often exhibiting change resulting from extension. To date, there is a conspicuous lack of in-depth study that seeks specifically to account for the domain-general cognitive mechanisms which underpin these analogical processes in the Insular Nordic context. In this connection, while the current chapter has provided a brief history of the respective inflectional systems of Insular Nordic, Chapter 3 elaborates on the cognitive underpinnings of analogical reasoning as a property of language use, of which language change is a function.

[14] Of course, it is not unlikely that doublets existed immediately after *u*-umlaut occurred in Ancient Nordic. However, the written record suggests that those in *a* had disappeared by the time that Insular Nordic is first attested.

3 Usage-based cognitive grammar and language change

The present chapter accounts for the chosen theoretical approach to the changes examined in Papers I–III. In 3.1, I compare and evaluate the applicability of two opposing theoretical approaches that seek to account for language change, namely the dual-processing model and the usage-based cognitive model. Section 3.2 fleshes out the usage-based cognitive approach through reference to rich memory for language (3.2.1), the impact of frequency on the structure of grammar (3.2.2), schematicity as a property of both linguistic and non-linguistic categories (3.2.3), and the prototype structure of categories as a function of cognitive economy (3.2.4). Subsection 3.2.5 provides a summary of Section 3.2. In Section 3.3, I elucidate the adaptive nature of categories as a function of analogy. Section 3.4 summarises the current chapter.

3.1 Two opposing theoretical approaches to language change

In this section, I provide an overview of two opposing theoretical approaches to language change, which, in the broadest terms, fall under usage-based cognitive grammar, on the one hand, and generative and dual-processing models, on the other. The aim of this endeavour is twofold:

1. To argue that language *use* is the mechanism of language *change*.
2. To demonstrate in light of 1. that a theoretical approach based on usage provides a more suitable framework to account for language change than one that does not.

The usage-based cognitive approach to language change is closely associated with the works of Jóhanna Barðdal (e.g. 2008), Joan Bybee (e.g. 1985, 1995, 2007, 2010, 2015), George Lakoff (1987), Ronald Langacker (e.g. 1987, 1988), and Michael Tomasello (e.g. 2000), to name but a few. Usage-based cognitive grammar places prominence on the impact of language use on transition and acquisition, the emergent structure of grammar, and language change. Therefore, in order to demonstrate the hand of language use in language change, it is first paramount to understand exactly what language use involves. Crucially, the process as conceived of here is not restricted to speaking and/or signing a language, but also includes

the tacit processes of receiving linguistic messages and decoding their content. As demonstrated with examples in what follows, this endeavour relies on prior experience with language that is stored in memory and available for the purpose of comparison, mapping, and extension of existing knowledge to new contexts.

Bybee (e.g. 2006) has characterised grammar as the cognitive organisation of the individual's experience with language. In the broader context of linguistic theory, the usage-based cognitive approach constitutes an abrupt departure from parsimonious, highly structural **dual-processing** model, itself an offshoot of **generative**/**transformational grammar**. The latter view economy as an inherent property of memory for language, by which redundant and predictable information is not stored due to constraints on "space" (e.g. Jakobson 1990:321; Rögnvaldsson 2013:139). In other words, the relevant experience is necessarily discarded from memory. In light of the economic view of memory for language, generative and dual-processing models seek to attribute linguistic ability to generative processes that are functions of genetic endowment, specifically evolved to facilitate (and exhibit) linguistic competence. According to generative grammar, the principles that guide competence constitute a biological faculty referred to as "Universal Grammar" (UG) (see e.g. Chomsky and Halle 1968; Clahsen 1999; Pinker 1999; also, more recently, Chomsky 2017 and Hauser et al. 2002).[15]

According to the approach from UG, the processes that facilitate language acquisition and linguistic performance are ascribed to evolutionary biology, by which the task of acquiring one's native language(s) is guided first and foremost by prescribed genetic constraints on the types of possible human language. In other words, proponents view the process as assisted by *innate* linguistic knowledge, an understanding of which is crucial in order to account for the principles that guide language acquisition and, more generally, the nature of human language (e.g. Chomsky 1965). However, due to the supposed innateness of UG, insight into its inherent linguistic principles can be blurred by various kinds of cognitive impairment. These include temporary states such as intoxication, momentary lapses in

[15] Despite the fact that generative and dual-processing models have evolved in terms of the instantiation-specific machinery applied in exposition, a fundamental adherence to parsimony has characterised the model since its inception. Indeed, a hyper-economical approach to memory for language is conspicuous in the *Distributed Morphology* (DM) paradigm, which assumes no lexicon, i.e. that part of memory supposedly responsible for storing necessary lexical information, such as lexical roots, unpredictable inflectional forms, and affixes. This tenet of DM has yielded the axiom "syntax all the way down", in allusion to the belief that fully formed inflectional forms result from concatenation of lexical heads (roots) and functional heads (morphemes = anything but a root) (e.g. Embick 2015; see Spencer 2019 for a detailed criticism of the theoretical foundations of DM espoused by Embick 2015).

retrieval that result in slips of the tongue, cognitive impairment due to stroke, to name but a few examples. For this reason, generative linguists consider language *use* a function of linguistic **performance**, i.e. the physical manifestation of language, an account of which pales in comparative theoretical importance to an understanding of linguistic **competence**, i.e. the system of grammar supposedly acquired with the assistance of UG (e.g. Chomsky 1965). Therefore, generative and dual-processing models view language use as an imperfect, or filtered-out, representation of linguistic competence.

The theoretical basis for envisaging linguistic competence as a function of genetic endowment is at best tentative. This is because the argument for UG is *a priori* in nature. In simple terms, UG was originally posited as part of a generative-driven endeavour to understand how human beings can know so much given their limited experience, a paradox characterised by Chomsky as 'Plato's Problem'. In linguistic-specific terms, the purpose of UG is to account for the acquisition of human language in spite of apparent significant inconsistencies in the linguistic input, i.e. Chomsky's *poverty of the stimulus* (e.g. Chomsky 1988:24). However, the so-called **language faculty**, i.e. the theorised biological host of UG, remains a contentious topic, both in linguistics and beyond (e.g. Pullum and Scholz 2002 for a critique of Chomsky's position).

Indeed, numerous scholars have argued against the existence of UG from fields such as evolutionary biology, linguistic typology, and psychology. For example, as part of an interdisciplinary analysis of language change in the evolutionary biological context, Croft (2002) characterises language as a metaphorical "moving target", whose structural properties change at such a rate as to evade eventual genetic encodement. In other words, language change occurs at too great a pace for what would be the feasible amount of time necessary to genetically encode any synchronic snapshot of linguistic competence. Further, according to the widely accepted axiom associated with Karl Popper (1959), a scientific theory is most appropriately defined as one that is falsifiable. In this connection, Dąbrowska (2015) points out that there is still no explicit evidence for UG, while Evans and Levinson (2009) have demonstrated that the differences between human languages permeate every level of linguistic function to so profound a degree as to render any theory of UG unfalsifiable.

In contradistinction to the innatism approach, usage-based cognitive linguists view linguistic competence as emergent from domain-general cognitive processes predicated on language use, the cognitive mechanism for which is the domain-general process of analogical reasoning (see e.g. Bybee 2010:6–8), defined as in (6).

(6) Analogical reasoning (analogy) is the cognitive process through which existing knowledge is extended to new contexts.

Inherent to the definition in (6) is that linguistic experience is stored in memory and provides the raw material for language use. In contradistinction to generative and dual-processing models, then, the approach from domain-general cognition assumes that, once stored, neither experience with language nor any resultant generalisations — also those characterised by redundant information — are discarded from memory but, rather, entrenched there for future reference (Bybee 2010:15; see Goldberg 2006:47–48).[16]

Further, language change provides empirical evidence for the impact of non-linguistic factors such as frequency and schematicity on language use and the direction of change. One common property of language change that elucidates this relation is the cross-linguistic tendency for highly frequent phrases to undergo phonetic reduction over time. Iterations of English *goodbye*, cf. *g'bye* and *bye*, all of which are incremental yields of gradual contraction from *God be with you*, provide just such examples. The meaning of this phrase is both highly predictable and discernible from context, meaning contraction is not an impediment to comprehension.[17]

[16] Charles Yang (e.g. 2016) espouses his *tolerance principle*, according to which high type frequency is the determinant of productivity, while token frequency, schematicity, and semantics are not attributed facilitating roles. In this connection, it is clear that knowledge of type frequency must be based on stored experience with language. However, one would also assume storage of knowledge of these other factors. One wonders, then, why the tolerance principle limits productivity to knowledge of type frequency, particularly when both token frequency and schematicity have been identified as contributing factors (see e.g. Axelsdóttir 2015, Jóhanna Barðdal 2006, and Markússon 2022a, 2022b, 2023a, 2023b). In other words, the *tolerance principle* seems easily disproved by observable productivity in actual language use and change (see Enger 2022 for a comprehensive criticism of Yang 2016).

The likely motivation for Yang's position is that, were productivity considered a function of token frequency, the model would necessarily have to assume storage of information pertaining to the frequency of occurrence of more than the bare minimum of "regular" forms. It has long been the belief in generative grammar and among proponents of dual-processing models for inflection that language users derive "regular" surface forms via the application of symbolic rules to a single lexical form, i.e. a stripped-down memory representation of a word that inflects according to a "regular" pattern (e.g. Rögnvaldsson 2013:137–139 on Icelandic). Yang is no exception in this regard. However, if information about the token frequency of "regular" forms is stored — remembering that "regular" is often conflated with "of high type frequency", which is in turn is often conflated with "productive", "predictable", and "no need for storage" (Herce 2019) — it would simultaneously be necessary to assume richer memory for language than proponents of generative and dual-processing models are typically comfortable with.

[17] See the discussion of the *reducing effect* in 3.2.1.

Conversely, high token frequency can facilitate the preservation of irregularity. For example, the infrequent English strong verbs *creep*, *leap*, and *weep* have developed regular past forms, cf. *creeped*, *leaped*, and *weeped*, beside older *crept*, *leapt*, *wept*, while highly frequent *keep*, *mean*, *sleep* alternate with past *kept*, *meant*, and *slept* only. This is considered a function of entrenchment and its impact on the availability of a linguistic item for use, i.e. that item's lexical strength.

It is, therefore, self-evident that if frequency impacts language use and change, knowledge pertaining to frequency is surely based on prior experience with language. Moreover, the relevant knowledge must be retained in memory, where language users have access to it. Indeed, if language users did not track frequency of use, frequency effects like those exemplified above should not be evident in language change. In other words, for prior experience with language to inform later usage choices — as it evidently does — memory of that experience must facilitate extension of the relevant knowledge to new contexts, cf. the definition of analogy in (6).

Despite clear evidence for language structure as emergent from the application of domain-general cognitive processes to existing knowledge, the approach from analogy has been challenged. In the historical context, this is perhaps understandable from the perspective of parsimony for mainly two reasons. First, as noted above, UG was posited as a working solution to a theoretical conundrum faced by linguists at a time when much less was known about the nature of linguistic storage. Secondly, the cognitive and articulatory apparatus is highly uniform across our species (Beckner et al. 2009), meaning that it is easy to find similarities across languages if that is what the researcher is looking for and then attribute similarity to UG. For this second reason in particular, proponents of UG have pointed to the apparent futility of recourse to analogy as an explanatory device when logically feasible analogies do not occur.

In this connection, Kiparsky (1974) has presented the explanation from analogy as theoretically wanting with his observation that conceivable analogies do not necessarily occur and are unlikely to do so in the future. In support of his point, Kiparsky offers the potential proportion in (7), which has yet not come to fruition.

(7) *ear* : *hear*
eye : X; X = **heye*

An initial criticism of the point that Kiparsky's equation is supposed to make is that proportional analogy is generally considered to occur between forms of words that belong to the same word class: *ear* is a noun, while *hear* is a verb. An obvious drawback of Kiparsky's observation is that Eng. *see* already exists, is highly frequent and, therefore, is unlikely to be forgotten and replaced (following e.g. Bybee 2015:95). Moreover, while it is true that the form **heye* is unlikely to be deduced from Eng. *eye* on the model of *ear* : *hear*, the irony of Kiparsky's point appears to escape him: the proportion in (7) constitutes a clear instance of analogy based on a linguistic model. Thus, the fact that the example may only have occurred in Kiparsky's mind by no means discounts recourse to analogy as the mechanism for linguistic innovation, even though it is not disseminated to the wider speech community.

It is highly probable that Kiparsky's observation follows in the spirit of King (1969:235), who characterised analogy as "a terminological receptacle devoid of explanatory power", a criticism levelled at the Neogrammarian habit of liberally ascribing all non-regular change to analogy. However, due to the very theoretical premise upon which the role of analogy is cast into doubt, criticisms inevitably prove to be strawmen. That is, the parsimonious nature of linguistic storage according to generative and dual-processing models has motivated attempts to reconcile synchrony with diachrony, by which change is characterised as changes to the system of **symbolic rules** that children supposedly acquire through intergenerational language transfer (e.g. Guðmundsdóttir 2008:27; Lightfoot 2006; see Beckner and Wedel 2010 for a critical view).[18] One posited function of this process was

[18] The application of symbolic rules to stripped down underlying lexical forms is a necessary means of "generating" the surface forms of "regularly" inflected words because, according to the generative model upon which Kiparsky's (1965) approach was based, only those forms of a word that are not fully predictable on the basis of other forms are stored in memory. By the same token, then, all predictable forms are derived from lexical form(s) by rule, thus sparing supposedly unnecessary strain on memory (e.g. Rögnvaldsson 2013:137–139; see above on the yet more stringent parsimony that is characteristic of DM).

An example of such an underlying form would be #barn-um#, from which the surface form Ice. dat.pl. *börnum* 'children' is supposedly derived through rounding of the root vowel *a* to *ö*, a process that some generative linguists have posited as synchronic *u*-umlaut (e.g. Þráinsson 2011, 2017; Rögnvaldsson 1981; see Markússon 2012, 2017 for a critical view). The underlying form #bar-num# is considered "stripped down" because it does not convey the rounded root vowel that is always present in the surface form. Such stripping down to the bare essentials is only necessary on a highly constrained view of the memory's capacity for linguistic storage. In the absence of such a view, it is not clear why a child acquiring Icelandic should derive a surface form from a stored form that it has never heard in the input. The question as to whether the input

grammar simplification, whereby the language learner chooses the simplest rule set based on the linguistic input (e.g. Kiparsky 1974). However, a highly valid critique of this position is that much language change does not simplify grammar.

For example, the innovative form Ice. fem. nom./acc.pl.**def**. *fætur**nar*** 'the feet' occasionally occurs instead of original masc. nom.pl.**def**. *fætur**nir*** and acc.pl.**def**. *fætur**na***. The fact that the feminine innovation eliminates the need for alternation between the latter pair of forms could of course be viewed as the simplification of grammar. However, such a view would be overly simplistic as use of fem. *fæturnar* redefines the typical dynamic such that overtly masculine nom.sg.**def**. *fótur**inn*** 'the foot' now alternates with overtly feminine *fætur**nar***. This can be considered a complicating development in light of the fact that singular forms of Ice. *fótur* almost never — if at all — occur as feminine (see PI:3 and sources cited there). In the broader context, it is an extremely rare relational dynamic within Icelandic noun paradigms that forms differ in grammatical gender (see PI:7 for an example). Therefore, it is doubtful that motivation for the deduction of fem. nom./acc.pl.def. *fæturnar* was the simplification of grammar in the broader sense (see Sims-Williams and Enger 2021 for further examples and discussion).[19]

Further, the claim that language change necessarily occurs during language acquisition is refuted by innovations such as Eng. past *creeped*, *leaped*, and *weeped*, which occur beside older past *crept*, *leapt*, and *wept* (see above). As noted by Bybee (2015:95), the verbs in question are unlikely to occur in child speech and, so, the weak innovations are attributable to adult language users (see Dąbrowska 2008, who shows that acquisition of Polish inflection is emergent and subject to reorganisation into adulthood). In light of these arguments, the generative approach to language change is considered insufficient to guide the analyses presented in Papers I–III, as it constrains the scope for permissible change by largely ignoring the impact of frequency, misunderstanding the motivations for change, and attributing change to the language acquisition phase only.

---

determines the shape of underlying forms or whether an underlying form determines the shape of the relevant surface form is a matter for discussion of the kind of logical fallacy inherent to circular reasoning.

[19] Kiparsky (2000) assesses analogical change in Gothic *ja*-stems and concludes that the relevant outcomes "do not complicate either the phonological or the morphological system of Gothic. The phonological constraints do not pick up any exceptions or morphological conditions. In particular, Sievers' Law ... continues to operate ... in a fully regular way." Thus, it appears that maintenance of or a move towards simplicity, characterised in terms of "grammar optimization", is viewed as validation in Kiparsky's approach to the nature and function of analogical change.

For the reasons stated above, I adopt a usage-based cognitive approach, which accounts for language change through reference to the application of domain-general cognitive processes to the linguistic input and stored memory representations for language across the language user's lifespan. This choice is motivated by four main factors. First, the approach has been somewhat lacking — particularly among Icelandic linguists concerned with synchronic variation, who in recent decades have typically followed the generative tradition (though see e.g. Axelsdóttir 2014, 2015; Ottósson 1992). Secondly, while traditional depictions of analogy suffice to delineate the outcome of innovation and change, these are often allowed to speak for themselves, while the domain-general cognitive specifics are merely implied, cf. proportions such as that in (7).

Thirdly, of all the so-called "domains", morphology poses the greatest challenge to the traditional generative view of productivity (see Bybee 1985:5–8). Indeed, the dual-processing model typically attributes a binary quality to productivity, i.e. linguistic structures are either regular, unmarked, and/or derived or irregular, marked, and/or stored (e.g. Clahsen 1999; Rögnvaldsson 2013; see Herce 2019 on the vagueness of such characterisations of "(ir)regularity"; also Langacker 1987 on what he characterises as the *rule/list fallacy*). The dual-processing view of productivity is, in turn, motivated by the model's parsimonious characterisation of linguistic storage, according to which productive patterns are extended through the application of symbolic rules in the absence of stored redundancies (see above). Conversely, non-productive patterns are assumed to proceed via direct retrieval from memory or via analogy (e.g. Clahsen 1999:996). However, as demonstrated by e.g. Barðdal (2008), productivity is a gradient property of language that is mediated by grammar-independent factors such as frequency and schematicity.

Fourth and finally, the impact of non-linguistic factors like frequency of use on usage choices, as evident through language change, is measurable. In other words, the usage-based approach enjoys support from beyond the realm of language, suggesting that it is falsifiable on terms that are not internal to the grammar nor the approach alone (see Bybee 2015:102). For the reasons just given, I consider both aims set out above to have been met. In support of adopting the usage-based cognitive approach, I elaborate on its main tenets in the next section.

3.2 The usage-based cognitive approach

The current section outlines the tenets, principles, and axioms of the usage-based cognitive approach to language change. In doing so, the current section also serves as a brief state of

the art via reference to notable and/or influential works within the field. The terminology associated with the usage-based cognitive approach will be defined through use of Insular Nordic examples where possible. Otherwise, examples are lifted from the works of others and acknowledged appropriately.

In 3.2.1, I argue in favour of rich memory representations for language based on observable frequency effects in language change. Subsection 3.2.2 accounts for the impact of frequency as a determinant of the structure of grammar. There, I first delineate the causal relation between frequency of occurrence, the domain-general process of entrenchment, and that property of memory termed *lexical strength*, before arguing that knowledge pertaining to frequency is a property of both specific and general experience. Subsection 3.2.3 discusses the interaction of schematicity with frequency as a determinant of productivity. In subsection 3.2.4, I delineate the *principle of cognitive economy*, i.e. the correlation between the hierarchical structure of categories and the breadth of applicability of different labels within the same category as a function of analogical reasoning. Subsection 3.2.5 summarises Section 3.2.

3.2.1 Rich memory for language

The current subsection elucidates the factors that support rich memory for language, picking up on arguments forwarded in 3.1 against the parsimonious view of storage associated with generative and dual-processing models. In usage-based specific terms, I identify three parameters for counting frequency — defined in terms of token frequency, type frequency, and (intra-paradigmatic) dispersion, utilising examples from both Insular Nordic and other languages. It is argued that the measurability of frequency renders the chosen approach testable and, therefore, scientifically viable as a theoretical framework employed to account for language change as a function of language use. Finally, examples are given of automatisation both in language and beyond, demonstrating that frequency provides a means of correlating practice with the overlap of articulatory and other motor gestures.

The usage-based cognitive approach to language contrasts starkly with the parsimonious view of memory characteristic of traditional generative and dual-processing models, as discussed in 3.1. Proponents of the latter assert that the capacity for linguistic memory is highly constrained and, therefore, characterised by minimal to no storage of predictable and/or redundant linguistic information. To resolve this issue, the generative and dual-processing position asserts that UG provides endowed mechanisms for linguistic competence that ensure minimal cognitive overload of memory for language (see e.g. Jakobson 1990:321). Further, the supposed means by which UG prevents overload is the

derivation of physical language or "surface forms" by the apparent application of symbolic rules to stripped-down underlying forms.[20]

Conversely, the usage-based cognitive approach assumes **rich memory for language**. In other words, the brain's capacity for storing linguistic experience is impressively large and no more constrained by the nature of such content than prevails for non-linguistic experience. Memories are strengthened by the repeated occurrence of "the same" action or event but may fade into obscurity if they are not enforced through continued experience (Bybee 2006:717–718). In contradistinction, traditional generative and dual-processing models posit an economically constrained lexicon, where information considered surplus to requirements for successful language processing is absent or, in some cases, even discarded from memory (see Bybee 2010:15; Langacker 1987). However, claiming rich memory for language is an empirically falsifiable proposition, as the impact that repeated use has on the strength of representations for language is tied to frequency of occurrence, which is in essence countable on the basis of recorded speech and text corpora (see below for examples). But what exactly is counted and how?

One method for counting frequency is defined in terms of **token frequency**, which refers to the number of times a linguistic item or phrase occurs in spoken language or text corpora. According to the usage-based cognitive view, the higher the token frequency associated with a linguistic item or phrase, the stronger its cognitive representation in memory. In the words of Bybee:

> If we metaphorically assume that a word can be written into the [mental] lexicon, then each time a word in processing is mapped onto its lexical representation it is as though the representation was traced over again, etching it with deeper and darker lines each time. Each time a word is heard and produced it leaves a slight trace in the lexicon, it increases in lexical strength. (Bybee 1985:117)

Hoffmann (2004), who in the main agrees with the view that frequency of use impacts the strength of memory representations for language, considers Bybee's characterisation of that impact too deterministic and, therefore, an oversimplification of the dynamics involved. Thus, Hoffmann (2004:189–192) argues for the interaction of frequency with semantic relatedness and **saliency**, characterised in terms of the preferred realisation of a given

[20] See footnote 18.

linguistic expression, cf. semantic doublets such as the complex prepositions Eng. *in light of* and *in **the** light of*, where the former is more frequent in American English; the latter in British English. Further, the preferred realisation may be more salient than the non-preferred realisation of another phrase even if the latter is of higher token frequency than the former, accounting for the entrenchment of some low frequency complex prepositions.

While I agree with Hoffmann's view, it by no means leads me to doubt the deterministic nature of the impact of frequency — all other things being equal. If a particular realisation of a linguistic item or phrase is more frequent than another realisation of that same item or phrase, the former will still be more strongly represented in memory. This view is evidenced by the often-noted correlation between the different token frequencies for inflectional forms of the same word and the direction of intra-paradigmatic levelling. In such cases, the prevalence of levelling in favour of the most frequent member of a paradigm demonstrates that although the same concept can have various realisations, i.e. distinct inflectional forms, the most frequent form has the strongest memory representation, i.e. is most *salient*, and is the most accessible for use. This property of storage is referred to as **lexical strength**.

Another method for counting frequency is defined in terms of **type frequency** and is determined by the number of individual items — each counted once — that conforms to a particular pattern. For this reason, type frequency can also be characterised as **dictionary frequency** (following e.g. Barðdal 2009:138), because it is the *existence* of a linguistic item that contributes to the type frequency of a pattern, rather than that item's token frequency. Thus, for example, the method for forming the past tense of English verbs such as *creep*, *keep*, *leap*, *mean*, *sleep*, and *weep* reflects a pattern of low type frequency because very few verbs inflect in this way, irrespective of each item's token frequency (see below).

Despite these parametric differences, both token and type frequency contribute to productivity (Barðdal 2006). As alluded to above, it is generally accepted in the usage-based cognitive literature that token frequency determines a linguistic item's lexical strength and, consequently, the ease with which an item can be accessed for use. Given this tenet of the approach, it is assumed that past forms such as Eng. *kept*, *meant*, and *slept* have significant lexical strength due to their high token frequency, while the past forms *crept*, *leapt*, and *wept* are relatively less well represented. In other words, the token frequency of the former set should correlate with relative ease of access, while the latter set are more likely to be forgotten — albeit only momentarily (see Bybee 2015:95). It is through such measurable

properties of use that token frequency and type frequency interact as testable determinants of change.

The pattern to which the English verbs mentioned above inflect is of low type frequency. In keeping with the correlation between frequency and lexical strength, the relevant pattern is less well represented than that represented by e.g. Eng. pres. *talk* ~ past *talked*, where the past stem is identical to that of the present, with the addition of *-ed*. There are literally thousands of English verbs that form their past tense in this way and, therefore, the usage-based cognitive approach assumes that this method both is well represented in memory and, resultantly, easily accessible at the cost of less common methods. This view is borne out through the interaction between token and type frequency: the less frequent past forms *crept*, *leapt*, and *wept* are occasionally forgotten and the most common method for forming the past tense is employed in such instances, cf. past *creep**ed**, leap**ed**, weep**ed***. Conversely, highly frequent past *kept*, *meant*, and *slept* have high lexical strength as a correlate of their token frequency and, therefore, are far less likely to undergo regularisation to past **keeped*, **meaned*, and **sleeped*.

Yet a third parameter for determining the impact of frequency involves establishing a linguistic item's **dispersion**, which Gries and Ellis (2015) define as the morphological contexts in which a linguistic exponent occurs. This property of use can be exemplified along *two* dimensions. For example, Paper I is concerned with the first dimension, where dispersion refers to the different rates at which the ending Ice. plural *-ur* occurs among feminine nouns, on the one hand, and masculines, on the other. The dispersion of the ending in question is highly skewed: Of the 14.92% of nouns in plural *-ur*, 91.89% of these are feminine; the rest masculine (see PI:2). Given the view that frequency of use impacts the strength of memory representations for language, we might assume that the dispersion of plural *-ur* should imbue the forms that contain it with **cue validity** for treatment as feminine rather than masculine — at least in momentary instances of uncertainty or when a new form containing the ending is encountered.[21] This hypothesis is tested in Paper I and shown to be correct.

The second dimension is characterised as **intra-paradigmatic dispersion** (Ice. *dreifitíðni*), which in Paper III refers to the number of paradigm cells that contains a particular stem variant. In this connection, consider the inflectional paradigm for OWN *fjǫrðr* 'fjord' in (8).

[21] Here and elsewhere, I follow Taylor's (2012:187) definition of cue validity: "The cue validity of feature *f* with respect to category C is the probability of C given *f*, i.e. $p(C \mid f)$."

(8) sg. nom. *fjǫrðr*
acc. *fjǫrð*
dat. *firði*
gen. *fjarðar*

pl. nom. *firðir*
acc. *fjǫrðu*
dat. *fjǫrðum*
gen. *fjarða*

It is determined that the dispersion of the stem variant *fjǫrð-*, which occurred in four cells of the paradigm, had double the rate of dispersion of the variant *firð-*, on the one hand, and the variant *fjarð-*, on the other: both of the latter were dispersed across two cells each. The question arises, then, as to whether the different rates of dispersion for individual stem variants facilitates association of a given variant with a particular morphological value associated with one or a set of paradigm cells. This line of inquiry informs part of the analysis in Paper III. Crucially, intra-paradigmatic dispersion may not contribute to lexical strength. Consider, for example, the fact that the dispersion of the stem variant *fjǫrð-* is twice as high as that of the variant *firð-*. However, as shown in Paper III (PIII:71), the latter variant is likely to have attained the greatest degree of entrenchment due to its occurrence in dat.sg. ***firði***, which is the most common form of *fjørður* as a constituent of composite place names.

Evidence for frequency as deterministic of entrenchment and lexical strength can be garnered from the performance of practiced behaviour — both linguistic and non-linguistic. A correlate of practice is commonly witnessed as different **frequency effects** in language change, specifically, the **reducing effect** and the **conserving effect**. The former is characterised by the gradual phonetic reduction of linguistic sequences — segments, affixes, words, sentences etc. — and is attributed to the process of **automatisation**, i.e. increased integration of repeated sequences of individual articulatory gestures. In other words, the more frequently a routine is practiced, the more integrated become the movements that make up that routine (see e.g. Langacker 2008:16; Bybee 2010:39–40 and sources cited there; also the discussion of Eng. *Goodbye* in 3.1).

An example of linguistic automatisation, characterised more specifically as **chunking**, can be lifted from Bybee (2010:41), who discusses the disparate degrees of gestural overlap between separate articulations of Eng. *I don't know*. Use of the phrase exhibits multiple

articulatory variants, cf. e.g. *I dunno*, and others that are even further contracted. By comparison, separate instances of *I don't inhale* exhibit little articulatory variation. Further, a highly contracted form of the latter would unlikely be understood, even when used in response to the question *Do you inhale?*, i.e. when use of *I don't inhale* might be expected. Similarly, semantically equivalent phrases in other languages undergo automatisation to respectively comparable degrees: Faroese *Eg veit tað ikki* 'I don't know' often takes the form [eɔiʰtʃː] or even [ɔiʰtʃː], while Far. *Ég innhaleri ikki* has no form that is contracted to any degree of equivalency.

The conserving effect has been described by Bybee and Thompson (1997:381), such that "...high frequency sequences ... resist change on the basis of newer productive patterns ..." This effect was exemplified above through reference to the development of the English strong verbs *keep*, *mean*, and *sleep*, on the one hand, and *creep*, *leap*, and *weep*, on the other. Due to their high frequency, individual forms of the former set are likely to have significant lexical strength. Conversely, forms of the latter set occur far less frequently and, as a direct consequence, are more likely to be forgotten — albeit momentarily.

The examples forwarded above provide direct evidence for rich memory representations for language. Further, as frequency of use can be correlated with specific frequency effects, this property of use also provides a measurable gauge for that correlation as a function of practice. Conversely, generative, and dual-processing models argue that linguistic units are either derived via the application of symbolic rules to stripped down, sometimes highly abstract, and preferably redundancy-free underlying forms ("regular" inflection) or retrieved directly from the lexicon ("irregular" inflection). According to the latter position, then, information about frequency of use is superfluous to requirements and, due to the economically constrained view of memory for language, not easily retained in memory.[22] Therefore, traditional generative approaches typically fail to account for the significantly varying degrees of automatisation exhibited among conventionalised linguistic sequences or resistance to regularisation as a correlate of frequency.

In order to successfully posit language use as an emergent function of domain-general cognition (see 3.1), it is necessary to demonstrate that the process is impacted by the same factors that facilitate automatisation in non-linguistic sequences. In this connection,

[22] This position is obviously the yield of circular reasoning: Does the nature of derivation from a single lexical form ease the burden of storage on memory — then as a function of genetic endowment — or does the economic nature of memory for language — then as a function of genetic endowment — determine the redundancy-free nature of derivation?

comparisons can be drawn with the repeated sequenced behaviour involved in the process of learning to use a keyboard or the performance of a complicated dance routine. Indeed, such processes are characterised by the performance of conventionalised sequences of individual actions that gradually exhibit increasing gestural overlap as a correlate of practice. The same cognitive and anatomical mechanisms that facilitate gestural overlap in non-linguistic sequences, such as dance routines, are considered the same as those that are at play in language use. In other words, the separate actions which constitute both linguistic and non-linguistic routines become conventionalised through gradual entrenchment of increasing gestural overlap due to repetition (Bybee 2010:34).

In terms of rich memory, then, prior experience — both linguistic and non-linguistic — must surely be stored as both specific and general memory representations (see Bod, Hay and Jannedy 2003; Bybee and McClelland 2005; Posner and Keele 1968). Otherwise, we should always be starting from the beginning each time a sequence of actions is repeated, with no stored basis for automatisation based on practice. Further, if prior experience was discarded from memory — a necessary entailment of dual-processing — we should also fail to break routines down into their individual constituent sequences once automated. Moreover, it is clear that frequency of use effects not only language change but, as part of that process, the structure of grammar also. For this reason, it is necessary to account for the relation between frequency and lexical strength, as well as the means by which both specific and general linguistic information comes to be stored as a function of rich memory. These properties of language use are discussed in the next subsection.

3.2.2 Usage and the structure of grammar

The current subsection states the core axiom of the usage-based cognitive approach, before discussing and subsequently demonstrating the impact of frequency on the domain-cognitive process of entrenchment as both a function and facilitator of analogy (as defined in 3.1). I then argue through reference to chunking that the impact of frequency on memory for language is driven by the implicit domain-general cognitive process of statistical learning. Finally, the discussion argues for statistical learning as a determinant of the domain-general process of categorisation and the hierarchical nature of taxonomies for varyingly specific/general schemas.

Subsection 3.2.1 argued for rich memory for language in light of observable frequency effects in language change. Indeed, without rich memory, we should not expect to observe correlations between change and frequency of occurrence. The examples forwarded in 3.2.1

are, therefore, taken to demonstrate that experience with language is not discarded from memory once the linguistic system is established but, rather, is stored and informs further usage choices, which pattern with prior experience. In this sense, grammar is emergent over the language user's lifetime (see 3.1). It is in this context that the core axiom of the usage-based cognitive approach to language change, stated in (9), proves highly relevant (following e.g. Bybee 2010:2; Bybee and McClelland 2005:382; Hay 2002; Hopper and Thompson 1984; Tomasello 2000).

(9) Language use determines the emergent structure of grammar, while that emergent structure determines further usage patterns.

Through explicit statement of this axiom, the usage-based cognitive approach proves free of the kind of (albeit inadvertent) circularity that characterises derivation from the underlying forms of generative grammar (see Bybee 2010:15; Langacker 1987; also footnote 18).

In order to account for the impact of frequency on grammar as emergent from use, as well as for the role of frequency in facilitating lexical strength, the domain-general cognitive process of entrenchment is here posited as a causative intermediary. In domain-general terms, Schmid (2017:3–4) defines **entrenchment** as the continual (re)organisation of knowledge in memory over the individual's lifetime, based on the mapping of new experience onto old. According to Gentner and Markman (1997), new experience is mapped onto stored experience(s) to the degree that the former reflects instances of the latter. In other words, the perception of similarity across instances of experience facilitates analogy, i.e. the process that extends existing knowledge to new contexts. As noted in 3.1, through reference to Bybee (1985), each time mapping occurs, it strengthens the representation of the relevant experience in memory.

Attributing the degree of entrenchment attained specifically to frequency of occurrence, Langacker asserts:

> Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact. With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit; moreover, units are variably entrenched depending on the frequency of their occurrence. (Langacker 1987:59)

In Langacker's view, entrenchment also diversifies variation exhibited across separate memory traces that stem from experience of "the same" or similar phenomena. This function stems from the fact that no two instances of what is otherwise considered "the same" experience are ever likely identical. In other words, mapping necessarily elaborates the emergent memory representation as a function of entrenchment. In turn, as the relevant knowledge is also emergent through entrenchment, diversification as a function of mapping continually increases the extensibility of that knowledge to new contexts. It is this function of rich memory that facilitates the storage of specific information. In turn, as a function of analogy, individual instances considered to represent the same or similar experience are subject to mapping between instances, yielding **schemas** of more general information, where 'schema' is defined as a superordinate concept that specifies the basic outline of multiple more specific concepts (following Tuggy 2007:83).

In their capacity as varyingly elaborate cognitive representations of experience, schemas serve as points of reference that are accessed for the purpose of categorising newly encountered phenomena based on similarity to prior experience (Taylor 2003:71). In the context of linguistic experience, then, the frequency with which a linguistic unit is used, i.e. the correlate of token frequency, or the number of items to which an inflectional pattern can be applied, i.e. the correlate of type frequency, determines the degree to which the relevant schema becomes entrenched. It is on the basis of this causal relation between frequency and entrenchment that a correlation between entrenchment and lexical strength can be discerned. Despite the opposing functions of the frequency effects exemplified in 3.2.1, the reducing effect should not be considered at odds with the claim that frequency strengthens cognitive representation. Indeed, as already noted, greater phonetic reduction is the correlate of practice. Automatisation — the mechanism of the reducing effect — often facilitates the process of **chunking**, by which sequences of distinct elements come to be analysed as gestalts (see Bybee 2010:34 and sources cited there). Langacker (1987; also Bybee 2010:44–46) argues that, as with all experience, the accessibility of such **chunks** from memory is dependent on the frequency of the relevant sequence as part of prior experience. Therefore, an inherent property of chunks is relatively deeper entrenchment in and greater accessibility from memory.[23]

[23] Note that greater accessibility is also a characteristic property of the *conserving effect* of frequency, demonstrating that distinct functions of frequency effects are by no means at odds.

Thus, it is possible to posit the causal relation *frequency* > *entrenchment* > *lexical strength*. Given the supporting arguments forwarded above, we should logically expect frequent linguistic units and patterns to be more accessible and, therefore, more frequently accessed, than those of lower frequency (see below). Indeed, Bybee (2015:95), in a similar spirit as Langacker, states that items and patterns of extreme low frequency are in danger of being forgotten and, eventually, disappearing, attributing atrophy to properties of lexical strength. In other words, when multiple linguistic items or patterns are in competition given the appropriate context of use, the most frequent option may usurp the "correct" form or pattern if the latter is temporarily forgotten due to low lexical strength (see the discussion of saliency in 3.2.1). Indeed, the process of levelling repeatedly demonstrates that the most frequent members of inflectional paradigms tend to win out in such circumstances, serving as the model for deduction of innovative forms within the paradigm (e.g. Bybee 1985, 2015:102; Tiersma 1982).

Based on the above, the impact of frequency on language structure is a measurable, falsifiable, and, therefore, scientifically sound basis upon which to account for the factors that determine the course of language change. Continuing in this spirit, then, the analyses presented in Papers I–III should support a view of language change as a function of the causal relation *frequency* > *entrenchment* > *lexical strength*. Now, it is necessary to delineate the mechanism for that impact. To this end, let us consider the question in (10).

(10) Do language users keep an exact tally of linguistic exemplars in memory or does such information accrue in the form of emergent schemas?

In order to answer this question, it is necessary to understand the domain-general process of **statistical learning**, through which tacit knowledge pertaining to frequency of use becomes entrenched in schematic form (see Jost and Christiansen 2017). In line with the discussion of both specific and general information, above, there is mounting evidence that language users retain memories of specific instances of language use (e.g. Bybee and McClelland 2005:384). However, as long since demonstrated by Posner and Keele (1968), schematic knowledge is accrued and elaborated on the basis of specific instances of prior experience that are considered the same or similar. In other words, the cognitive representation of general experience is predicated on prior mental organisation of specific information (Gerken 2010).

Rosch (e.g. 1975, 1978) has demonstrated that **categorisation** is the domain-general process through which living organisms cognitively organise their environment by assigning structure to the phenomena they encounter in it. The process of categorisation, therefore, serves as a means of **uncertainty reduction**.[24] However, Rosch et al. (1976) show through a series of experiments that categories are by no means defined in terms of the necessary and sufficient features that characterise the classical approach to categorisation: the notion that in order to belong to a category, a given phenomenon must necessarily exhibit a specific set of features and that — in circular fashion — if the necessary features are discernible in one and the same entity, these are sufficient for assignment to the relevant category (see Taylor 2012:186).

Rather, Rosch and colleagues demonstrated that categories can be elaborated to such a level of abstraction as to contain phenomena which at lower levels would represent distinct, i.e. respectively specific, instantiations of individual category members. In other words, thinkers are able to generalise over instances of experience with specific items considered the same or similar so that a more general schema is elaborated. Crucially, both specific and general information, i.e. schematic knowledge representative of various levels of abstraction, are retained in memory and form hierarchical taxonomies of varyingly abstract schemas. For example, Rosch et al. argue that the category FURNITURE is more abstract than hierarchically subordinate categories such as CHAIR, TABLE, BED. In turn, CHAIR is more abstract than e.g. DINING CHAIR, ARMCHAIR, BEANBAG; TABLE is more abstract than COFFEE TABLE, DINING ROOM TABLE, DESK; BED is more abstract than DOUBLE BED, SOFA BED, HAMMOCK.

For examples of such taxonomic structure within linguistic categories, consider alternation of the kind OWN/Ice. masc. nom.pl. *-ar* ~ acc.pl. *-a* and nom.pl. *-ir* ~ acc.pl. *-i*, cf. *fugl**ar*** ~ *fugl**a*** ‘birds’ and *gest**ir*** ~ *gest**i*** ‘birds’. While alternation of this kind has been levelled in Faroese, cf. Far. nom./acc.pl. *fuglar, gestir,* Icelandic has both retained and extended both patterns to new paradigms, cf. the loans Ice. nom.pl. *iPad**ar*** ~ acc.pl. *iPad**a*** ‘iPads’, nom.pl. *bar**ir*** ~ acc.pl. *bar**i*** ‘bars’. Crucial in the context of the current discussion is that extension never yields a mixed pattern: alternation of the kind **nom.pl. *-ar* ~ acc.pl. *-i* or ***-ir* ~ *-a* never occurs.

[24] Uncertainty in the context of inflectional morphology is often defined and quantified in terms of **entropy**, i.e. the degree of uncertainty that pertains to the predictability of one inflectional form on the basis of others from the same paradigm (see e.g. Cser 2023, who quantifies the entropy of inflectional forms within regular Latin verb paradigms).

Relating this to the emergence of general information from specific information, the fact that these functionally identical yet formally distinct relations are systematically extended to new contexts in Icelandic suggests that language users discern a partial identity relation between the relevant forms *within* each class. However, it can be argued that, simultaneously, yet more schematic knowledge precludes mixing of the respective inflectional patterns. In other words, while application of each *specific* pattern proceeds on a word-by-word basis, the *general* knowledge that the individual patterns are distinct necessarily represents a greater level of abstraction. To demonstrate, consider the "kind of" rule of referral for Icelandic stated in (11).[25]

(11) If the nominative plural of a masculine noun ends in *-ar*, its accusative plural form will end in *-a* (and vice versa).

As well as that stated in (12).

(12) If the nominative plural of a masculine noun ends in *-ir*, its accusative plural form will end in *-i* (and vice versa).

The stipulations in (11) and (12) imply more abstract relational knowledge that can be represented by the **sister schemas** $[\text{-V}_1r]_{\text{nom.pl.}} \sim [\text{-V}_1]_{\text{acc.pl.}}$, where subscript '$_1$' indicates phonetic identity between intra-paradigmatic instances of V, i.e. {a, i}.[26] Indeed, were more abstract relational knowledge not at play, use of the relevant forms should not preclude alternation of the kind **nom.pl. *-ar* ~ acc.pl. *-i* or **-*ir* ~ *-a* to the degree that actual usage patterns show. Now, compare (11) and (12) with the statement in (13).

[25] Hansson (2007:92) defines rules of referral as stipulations that account for syncretism within an inflectional paradigm. The "rule" states a purely morphological fact, whereby "the realization of one inflectional form is systematically 'referred to' that of another form within the same paradigm". Above, "kind of" refers to the near-syncretism stipulated by the "rules" in (11) and (12). Rules of referral are thematically akin to e.g. Blevins's (2016:105) idea of "implication as uncertainty reduction" and Wurzel's (1984:208) *implicative paradigm structure conditions* (PSCs) and are not to be confused with the symbolic rules of generative grammar.

[26] Audring (2019) defines the relation between sister schemas in terms of equivalency in the level of complexity conveyed. Such schemas are referenced as a means of checking pertinent semantic and/or formal distinctions between the sets of forms over which the relation abstracts.

(13) Alternation of the kind nom.pl. -V_1*r* ~ acc.pl. -V_2 (V = {a, i}) never occurs in nouns.

The statement in (13) is necessarily more abstract and less conditional than those in (11) and (12) — even when the latter are stated in the most schematic terms, i.e. through reference to the sister schemas $[-V_1r]_{\text{nom.pl.}} \sim [-V_1]_{\text{acc.pl.}}$. This is because (11) and (12) must first stipulate the kinds of alternation that manifest the relevant schema.

The immediately preceding preamble can also be framed in terms of frequency of use and the structure of grammar, as the respective patterns are extended at different rates. Thus, alternation of the kind nom.pl. *-ar* ~ acc.pl. *-a* has been far more productive through the centuries than nom.pl. *-ir* ~ acc.pl. *-i*. In this connection, Bybee (e.g. 2010:67) notes a correlation between high type frequency and productivity. Crucially, the vast majority of Icelandic masculine nouns exhibits the pattern nom.pl. *-ar* ~ acc.pl. *-a*, while those that have nom.pl. *-ir* ~ acc.pl. *-i* are relatively few (Svavarsdóttir 1993:105). Therefore, their different rates of productivity should come as no surprise, but clearly support a correlation of the productivity that a given pattern exhibits and the type frequency of that pattern.

The preceding discussion necessitates the question: How can language users acquire, retain, and utilise knowledge pertaining to frequency unless they keep count of each instance of use? The answer is: They cannot. However, the frequency with which a pattern's schema is referenced and instantiated continually entrenches it in memory (e.g. Cordes 2017; Janda 2002). In other words, the process of statistical learning occurs tacitly, continuously reshaping schematically represented knowledge pertaining to frequency of occurrence. As the relevant knowledge becomes further entrenched, schemas are elaborated to accommodate new instances of experience. This process by no means requires that existing knowledge or less abstract schemas be discarded from memory in the fashion of more parsimonious approaches to linguistic storage (see 3.1).

It is, therefore, unnecessary to assume that specific memories of every single instance of use are readily accessible from memory, even though some may well be (see above). In other words, language users may categorise via reference to a specific memory, although, as experience accrues, categorisation is increasingly likely to proceed via reference to schemas whose inherent generalisations are informed by the application of statistical learning to prior experience. Further, by its nature, the same process teaches the language user to rely on the cue validity of one inflectional form for the intra-paradigmatic occurrence of another (see 3.2.1). Thus, users of Icelandic know that nom.pl. *-ar* implies acc.pl. *-a* (and vice versa)

because alternation of the kind **nom.pl. *-ar* ~ acc.pl. *-i* or ***-ir* ~ *-a* does not occur and, therefore, does not inform the entrenchment of the relevant schemas.[27]

With regard to the structure of grammar, it is argued above that this is emergent from the impact of frequency of use on domain-general processes such as statistical learning, entrenchment, generalisation, and categorisation. Thus, in Hoffmann's (2004) terminology, the most frequent instantiations of a concept or pattern are the most deeply entrenched in memory and, therefore, the most *salient* also. Due to the causal relation *frequency* > *entrenchment* > *lexical strength*, the relevant schemas for deeply entrenched knowledge are more readily accessible and, therefore, more likely to be applied as a means of uncertainty reduction when categorising previously unencountered phenomena, including the application of inflectional patterns to borrowings, as exemplified above. This then begs the question: Are the factors that facilitate entrenchment strictly deterministic? Providing an answer to this question is the aim of the next subsection.

3.2.3 Schematicity and its interaction with frequency

In the preceding subsection, I argued for the causal relation *frequency* > *entrenchment* > *lexical strength* and offered examples from Icelandic which indicate that highly frequent patterns are more likely than less frequent patterns to be applied in the inflection of borrowings. As argued in 3.2.2, this property of usage impacts the structure of grammar. However, it is also necessary to account for those instances in which less frequent patterns are extended. To achieve this aim, I discuss here another non-linguistic factor that drives language change, referred to as **schematicity** and defined as the degree of phonetic dissimilarity among the members of a class (2010:67). As demonstrated in Papers I and II, the impact of minimal schematicity is most obviously viewed as limited productivity, i.e. a relatively low rate at which a pattern is extended to new contexts (see below). The interaction of schematicity with frequency as a determinant of productivity is the focus of the current subsection.

As noted in 3.2.3, high schematicity and high type frequency can interact to facilitate all but unlimited productivity. The class of English verbs in past *-ed* can be used to exemplify the combined effects of these factors. The class in question contains literally thousands of verbs, the numbers of which are constantly increasing. This means that entrenchment of the relevant pattern is continually enforced and, therefore, ever more easily accessed for

[27] See Björnsdóttir (2021, 2023) on the nature of generalizations in grammatical gender and inflectional morphology for Icelandic nouns from a child learning perspective.

application. Further, the class lacks clear phonetic definition, meaning that, as well as being the pattern with the strongest representation in memory, it is also applicable to items of basically any shape.

Despite the combined effects just discussed, schematicity can be a facilitating factor in productivity in its own right, i.e. in the absence of significant type frequency. Thus, patterns of medium to low type frequency can exhibit limited productivity if members of the relevant class are phonetically and/or semantically similar, i.e. exhibit low schematicity. In this connection, Barðdal (2008:1) characterises the productivity of syntactic construction in terms of their attractive force. Thus, a construction is able to attract items that are similar to others that already fill its open slot(s). Further, Barðdal (2008:9) posits an inverse relation between type frequency and semantic coherence which, she argues, adequately predicts the varying degrees of productivity exhibited by different constructions. On the limited productivity of inflection classes, Bybee (e.g. 2010:69) attributes the mechanism for productivity among low-frequency, low-schematicity classes to the **gang effect**, by which a high concentration of common phonetic properties among a small class of items suffices to attract new members.

An example of the gang effect is betrayed by the development of the formerly weak English verb *wear*, which now has the standard past form *wore*. Framing deduction of the past form in the context of interaction between frequency and schematicity, it is noteworthy that the relevant pattern has highly limited productivity. Thus, it originally applied to three verbs only, i.e. *bear*, *swear*, and *tear*, cf. past *bore*, *swore*, and *tore*. Due to the deterministic nature of the causal relation *frequency* > *entrenchment* > *lexical strength*, argued for in 3.2.2, representation of the relevant pattern in memory was likely weaker than that of many other methods of forming the past tense — especially relative to that which adds *-ed* to the present stem. We should, therefore, expect the pattern $[X\textit{ear}]_{\text{pres.}} \sim [X\textit{ore}]_{\text{past}}$ to be less accessible for use than many others — all other things being equal.

Further, the class containing *bear*, *swear*, and *tear* was minimally schematic, making it unlikely to attract verbs — weak or strong — of different phonetic shapes, such as Eng. *go*, *forget*, *talk*, *write*, which are phonetically dissimilar both from each other and from members of the *bear ~ bore* class. Nonetheless, the *bear ~ bore* class was able to attract weak *wear* on the basis of phonetic coherence with existing members. Thus, while frequency is a positively deterministic factor in language change, the preclusive impact of extreme low frequency on productivity can be counteracted by minimal to low schematicity in order to facilitate limited productivity by means of the gang effect. This is demonstrated by the examples forwarded above and in the analyses presented in Papers I and II.

3.2.4 Categories and the principle of cognitive economy

The current subsection fleshes out the relation between the taxonomic structure of categories and the applicability of category labels in categorisation, in accordance with a principle of cognitive economy. To this end, I delineate the taxonomic relation between varyingly schematic categories for formally and/or functionally related items and demonstrate the level of abstraction at which a category is most extensible to potential inductees. The level of abstraction that accommodates the most broadly applicable category labels is identified as the level that most suitably applies in neutral reference, constituting the basic level of a three-tier hierarchy.

Through a series of experiments, Rosch (1975; see also Rosch et al. 1976) argued in favour of a **principle of cognitive economy**, by which *basic* category labels and contents convey maximum information with optimally minimal cognitive cost (see below). In order to demonstrate the domain-general nature of this principle, distinct non-linguistic category labels applied at respective levels of abstraction within functionally continuous taxonomies can first serve as exemplary. For this purpose, I will utilise the category FURNITURE, the taxonomic structure of which was delineated in 3.2.2, showing the applicability of that category label to members represented at different levels of abstraction.

Thus, the category FURNITURE betrays abstraction to a high degree of schematicity. This is because no single feature or sets of features suffice to represent all individual category members with minimal cognitive cost. However, at the medially schematic level, features are both general enough as to motivate neutral reference to individual members but specific enough to distinguish them from those of another category label that is representative of the same level of abstraction. Therefore, CHAIR, BED, and TABLE serve to distinguish individual types of the highly schematic category FURNITURE, both from each other and from members of the category DOG, which is a medially schematic subcategory of highly abstract (NON-HUMAN) ANIMAL.

In turn, at the minimally schematic level, category labels such as DINING CHAIR, ARMCHAIR, and GARDEN CHAIR distinguish members of the medially schematic category CHAIR; COFFEE TABLE, DINING ROOM TABLE, DESK distinguish members of the medially schematic category TABLE; DOUBLE BED, SOFA BED, CAMPER BED distinguish members of the medially schematic category BED. In light of the representational scope of the respective category labels, the categories CHAIR, BED, and TABLE are considered **basic level categories**. Thus, the category CHAIR is not schematic enough to apply to items such as double beds or dining room tables but nor is it too specific as to encompass dining room chairs only; garden

chairs, rocking chairs, and armchairs can also be assigned the category label CHAIR. In this sense, CHAIR, BED, and TABLE occupy "[t]he level used for everyday neutral reference", a property that Croft and Cruse (2004:83) attribute to basic level categories.

The less schematic status of e.g. GARDEN CHAIR and ROCKING CHAIR within the taxonomy dominated by the highly schematic category FURNITURE accounts, then, for their function as single points of reference for categorisation of chairs that are otherwise formally and functionally dissimilar. By the same token, highly schematic FURNITURE is considered **superordinate** to basic level CHAIR, as the former has "…fewer defining attributes than…" the latter (Croft and Cruse 2004:84). Finally, category labels such as GARDEN CHAIR and ROCKING CHAIR reflect **subordinate level categories**, characterised as those in which "…members have high mutual resemblance…" (Croft and Cruse 2004:85). Perception of similarity to the members of minimally schematic categories, i.e. those that represent the subordinate level, should provide for the highest number of potential one-to-one formal correspondences across sets of functionally related items.

Crucial to Rosch's characterisation of categories is that they exhibit **prototype structure**, i.e. some items are considered more prototypical of the category as a whole than others. In this sense, a dining chair might be considered more "chair-like" than an armchair, while the latter might in turn be considered more prototypical than a beanbag (see Rosch 1975 for experimental evidence in support of prototype structure). Thus, the common attributes of e.g. chairs are not determined by necessary and sufficient features (see 3.2.2), but, rather, converge upon the most frequent formal and functional attributes common to existing category members. In turn, the schema that abstracts over the formal and functional attributes common to all chairs informs not future assignment of items to the category, but also the status of those items with respect to prototypicality. Thus, prototypicality is best characterised in terms of **family resemblance**, by which category members share core attributes, but may differ both in form and the means by which they serve their intended function.

The principle of cognitive economy is also a property of the structure of linguistic categories. Accordingly, schemas for such categories arrange themselves in hierarchic taxonomies, where category labels represent varyingly specific and general degrees of abstraction. Further, linguistic categories exhibit family resemblance structure that centres around prototypes. Moreover, prototype effects are observable through the varying rates at which different items are attracted by a given constituent schema of the taxonomy. This property of linguistic categories is discussed in the next section, and in more detail in Section 3.3.

3.2.5 Section summary

The current section has argued for the impact of domain-general cognitive processes on language use, which is viewed as the mechanism of language change. Subsection 3.2.1 argued in support of rich memory for language, identifying some of the different types of frequency effects that are observable through language change. Section 3.2.2 discussed frequency as a determinant of the structure of grammar, arguing for the causal relation *frequency* > *entrenchment* > *lexical strength*, before accounting for the tacit accrual of knowledge pertaining to frequency via statistical learning. Section 3.2.3 discussed the interaction between schematicity and frequency and its impact on the direction of change.

Subsection, 3.2.4 delineated the cognitive means by which items are assigned to functionally related categories that constitute different levels of a three-tier taxonomy, defined by increasing/decreasing schematicity. The basic level, i.e. medially schematic categories, are considered those that best encompass all functionally identical but formally different category members and, therefore, is used for everyday neutral reference. Finally, it was argued that the category-internal taxonomies that are a function of the principle of cognitive economy entail prototype structure, by which some members are considered better representatives of the category, as these share different iterations of formal and functional attributes with all other category members. Thus, the structure of categories is characterised by family resemblance, rather than specific sets of necessary and sufficient features.

3.3 Analogy and the adaptive nature of categories

In the current section, I posit analogy as the mechanism of domain-general cognitive processes applied to both linguistic and non-linguistic tasks. Before the discussion commences, reiteration of the usage-based cognitive approach's core axiom — repeated in (14), below, from (9) — seems pertinent.

(14) Language use determines the emergent structure of grammar, while that emergent structure determines further usage patterns.

So, the cycle of use continues as a function of domain-general human cognition (e.g. Beckner et al. 2009).

In terms of the cognitive mechanism of language use, the current thesis follows Blevins and Blevins (2009:1), who observe that "[t]here is mounting evidence from work in cognitive psychology that the talent for analogical reasoning constitutes the core of human cognition"

(citing Penn et al. 2008), and that "analogy may be a highly domain-independent cognitive process" (citing Halford and Andrews 2007). In what follows, analogy is considered one fundamental mechanism of language change, which, in turn, is viewed as a function of language use.[28]

Subsection 1.1 alluded to two distinct functions of analogy: extension and levelling. In the context of the current thesis, the former process is defined as the application of an existing *pattern* of intra-paradigmatic alternation in a paradigm that formerly did not show it. The latter process is characterised as the deduction of an innovative inflectional form containing a stem variant previously found in another cell or cells of the paradigm (see below for examples). Thus, strictly speaking, extension involves the influence of external patterns on the internal relational dynamics of paradigms, while levelling occurs within one and the same paradigm (though see below on functional crossover between the two processes).

Examples of analogical extension can be taken from Papers I and II. Thus, for example, Paper I posits the influence of feminine nouns in plural *-ur* on masculines in the same ending, by which the latter are reanalysed as grammatically feminine due to the ending's 91.89% rate of dispersion among feminine nouns. Paper I argues that, on account of this property of use, masculine forms such as plural *fætur* 'feet' sometimes alternate with overtly feminine forms, cf. fem.pl.**def**. *fætur**nar*** instead of masc.nom.pl.**def**. *fætur**nir***, masc.acc.pl.**def**. *fætur**na***. The model for this pattern of alternation is exemplified by relations such as fem.pl. *stelpur* 'girls' ~ fem.pl.**def**. *stelpur**nar***. Extension is considered the archetypal manifestation of **proportional analogy**, as depicted by the proportion in (15).

(15) pl. *stelpur* : fem.pl.def. *stelpurnar*
pl. *fætur* : fem.pl.def. X; X = *fæturnar*

The function of proportions can be delineated in terms of the schema A : B :: C : D (see e.g. Fertig 2013:28). Thus, (15) suggests that the pattern of alternation between the A- and B-forms, i.e. *stelpur* and *stelpurnar*, respectively, is extended to the relevant cells of the paradigm for Ice. *fótur* 'foot', such that the functionally equivalent C- and D-forms, i.e. *fætur*

[28] Bybee and Beckner (2015:503) assert that categorisation is the most fundamental cognitive process in language change. However, as noted by Rosch (e.g. 1975), categorisation is dependent on the perception of common attributes. Further, as Gentner and Markman (1997:48) have argued, the perception of *functional* similarity is a prerequisite for analogical reasoning. Therefore, the position is taken here that categorisation cannot proceed unless the newly encountered phenomenon exhibits similarity to another that forms part of prior experience. In other words, categorisation is here considered a function of analogy (see below).

and X, come to exhibit the pattern modelled by A and B. While the **source** of the C ~ D-relation is the C-form, it is formal similarity and functional equivalence between A and C that motivates the extension of the A ~ B relation to C and D, hence X = *fœturnar*, i.e. the **target** D-form.

Similarly, in levelling, a target comes to occur in the same paradigm as the source upon which the former is based. An example of levelling taken from Paper III is the extension of the stem variant *firð-* in the paradigm of Far. *fjørður* to the dative plural cell, cf. innovative dat.pl. ***firð****um*, beside older dat.pl. *fjørðum*. Prior to levelling, the variant *firð-* occurred in the cells of the dative singular, cf. dat.sg. ***firð****i*, and the nominative/accusative plural only, cf. nom./acc.pl. ***firð****ir*.[29] However, as noted by Bybee (2015:95), levelling does not involve the transformation of an older form into a younger one. In this connection, it is important to note that older dat.pl. *fjørðum* still exists alongside its younger cellmate.

However, despite the fact that paradigm-internal constraints determine the source of levelling, functional motivation for the process can also be characterised as proportional. In brief, Paper III posits dat.sg. *firði* as the source of levelling to dat.pl. ***firð****um*. Crucially, in Faroese, it is most common that all forms of a given paradigm contain the same stem variant. Indeed, widespread levelling of older *ø* to *a* in dative plural stems is suggestive of a general tendency to level vocalic alternation, cf. e.g. nom.sg. *hv****a****lur*, acc.sg. *hv****a****l*, dat.sg. *hv****a****li*, nom./acc.pl. *hv****a****lir*, dat.pl. *hv****ø****lum*/innovative *hv****a****lum* (Þráinsson et al. 2012:398–399). Thus, the proportion in (16) demonstrates that external influence is also a motivating factor in levelling.[30]

(16) dat.sg. *hva*$_1$*li* : dat.pl. *hva*$_1$*lum*
dat.sg. *fi*$_2$*rði* : X; X = *fi*$_2$*rðum*

[29] It is necessary to note that the stem variant *firð-* has also found its way to the cell of the accusative plural, cf. OWN *fjǫrðu*, Far. *firðir*. However, this process is most likely the result of levelling that occurred in all masculine paradigms where the plural nominative and accusative forms were once formally distinct (see PIII:73). Therefore, the change in question can be attributed to the impact of type frequency on a smaller inflection class, rather than that of the token frequency of dat.sg. *firði*.

[30] Subscript '$_1$' and '$_2$' in (16) are intended to draw attention to phonetic identity between stem vowels of distinct forms within the respective paradigms.

Although the proportion does not specifically state it, (16) is representative of the fact that levelling does not lead to change that is incongruent with existing patterns.[31] Therefore, in light of the preceding discussion, the definition of analogy in (6) can be elaborated as in (17) so as to encompass both extension and levelling (following e.g. Gentner 2005; Christie and Gentner 2010).

(17) Analogy is the extension of existing *relational* knowledge to new contexts.

In (17), "relational" implies that argument structure is aligned across scenarios that are perceived as functionally equivalent, as delineated below with both linguistic and non-linguistic examples.[32] However, in order to better grasp what the alignment of argument structure entails, it is first necessary to understand the functions of analogy that motivate it. In this connection, Gentner and Markman (1997) have shown that analogy facilitates categorisation, a prerequisite of which is structure-mapping between the relations that cause a set of scenarios to be perceived as similar (see discussion in 3.2.2). Perception of relational similarity is optimised by the identification of **parallel connectivity**, i.e. the state where the items that constitute individual arguments *within* separate relations correspond on a one-to-one basis according to role (e.g. Gentner and Hoyos 2017:675).

To take an example from Kotovsky and Gentner (1996), the argument structure involved in the action of a toddler waving its arm and causing a cup to fall from a table can be aligned with that of a squirrel swishing its tail and causing an acorn to fall from a tree. Thus, the toddler aligns with the squirrel; the cup with the acorn; and the table with the tree on a one-to-one basis. Crucially, there is little formal similarity between a toddler and a squirrel; a cup and an acorn; a table and a tree. However, similarity *among* relations is dependent on the perception of common *relational* structure across a set of scenarios and, therefore, formal differences between items do little to interfere with that perception (e.g.

[31] Of course, the process of levelling has taken various directions within the paradigm of Far. *fjørður*, leading to a unique inflectional dynamic among Faroese nouns (see Sections 6.2.1 and 6.3.2). However, it is considered likely that each instance of levelling accorded with an existing pattern, although the sum of different analogical operation has blurred the initial intent for each individual development.

[32] See Bulloch and Opfer (2009) for experimental findings that demonstrate the alignment of argument structure across non-linguistic domains among both children and adults.

Christie and Gentner 2007; Gentner 1983, 2005; Kotovsky and Gentner 1996; Krawczyk, Holyoak, and Hummel 2004, 2005; Markman 1997; Markman and Stilwell 2001).[33]

Scholars have demonstrated repeatedly that one-to-one mapping between intrinsic formal and/or functional attributes facilitates elaboration of a category's existing schema (see Christie and Gentner 2010; Cordes 2017; Gentner 2005; Gentner and Hoyos 2017; Kotovsky and Gentner 1996). If a new inductee to a category exhibits a degree of formal dissimilarity from existing members, yet all share a functional attribute or set of functional attributes, further abstraction will ensure that the relevant schema is "fully compatible with all the members of the category it defines…" (Langacker 1987:371). Accordingly, the toddler-cup-table and squirrel-acorn-tree scenarios mentioned above could well be assigned to a category such as MOTIONS THAT MAKE OBJECTS FALL FROM SURFACES if they constitute a part of prior experience.[34]

Thus, as the process of schematisation abstracts further away from all but the common attributes of the category, its schematicity increases, permitting a greater degree of dissimilarity among both current and future category members, all of which nonetheless comply with the evermore abstract, emerging schema. The dynamics that prevail between the different members of a category as a function of schematisation have been characterised by Wittgenstein (1978) in terms of **family resemblance**. The category PREDATOR provides an example of a highly schematic *relational* category whose members exhibit family resemblance structure as a property of parallel connectivity. The common attributes of predators are necessarily predicated on *functional* opposition to prey, rather than alignment of *formal* attributes common to predators. In other words, while e.g. polar bears and Venus flytraps have little in common in terms of appearance — indeed, they do not even belong to the same biological kingdom — both are predatory species. Likewise, there is little formal

[33] Gentner and Markman (1997) demonstrate with the proportion 1 : 3 :: 3 : 9 that perception of formal similarity is unnecessary for a successful analogy. Thus, despite the fact that 1 and 3, i.e. the A- and C-forms of the proportion, and 3 and 9, i.e. the B- and D-forms, have nothing in common in terms of form, pronunciation, or meaning, it is the *relation* between the exemplary A- and B-forms that facilitates analogy, i.e. functionally equivalent items of the proportion exhibit parallel connectivity and can, therefore, be aligned according to role: 1 goes thrice into 3 and, on that basis, the relation between 3 and 9 is understood as functionally equivalent. This is not to say that intrinsic formal attributes, common to a set of items, *never* assist in establishing parallel connectivity. Indeed, the greater the degree of object similarity perceived between a set of items, the greater the likelihood that their common attributes will motivate comparison, structure-mapping and, ultimately, facilitate alignment.

[34] See e.g. Lakoff (2018:43) on categories established "on the fly", such as THINGS YOU TAKE ON A PICNIC.

identity between seals and flies, although both of these can be aligned as prey according to the relationship of each with their respective predator types.[35]

Gentner characterises relational categories as those:

> ...whose meanings consist either of (a) relations with other entities, as in *predator* or *gift*, or (b) internal relations among a set of components, as in *robbery* or *central force system*.
> (Gentner 2005:245)

She continues:

> Relational categories contrast with object categories ... Of course, object categories contain not only property information but also relational information. For example, that tigers hunt and eat animals is part of our concept of a tiger, along with intrinsic attributes such as their stripes. (Gentner 2005:245)

Returning to the significance of relational knowledge for the definition of analogy presented in (17), the degree to which common relational structure can be perceived determines the likelihood that existing relational knowledge is extended to new contexts, simultaneously elaborating the relevant schema. The view of categorisation just outlined can now be delineated with linguistic examples, which can be picked up from the discussion of Old West Nordic and Icelandic masculine nouns in nom.pl. *-ar* ~ acc.pl. *-a*, e.g. OWN/Ice. *fugl**ar*** ~ *fugl**a***, cf. borrowed *iPad**ar*** ~ *iPad**a***, and those in nom.pl. *-ir* ~ acc.pl. *-i*, e.g. *gest**ir*** ~ *gest**i***, cf. borrowed *bar**ir*** ~ *bar**i*** (see 3.2.2).

In terms of common formal attributes, it is clear that extension of the relevant patterns to the paradigms of borrowed masc. *iPad* and *bar* is minimally dependent on formal similarity. Indeed, *fugl* and *iPad*, on the one hand, and *gestur* and *bar*, on the other, have little to nothing in common in terms of phonetic structure. However, the borrowings have been

[35] Note also that the category has been expanded to include human beings on metaphorical grounds, as exemplified by its application to people — particularly men — who take advantage of other people — particularly women — for sexual gratification. Likewise, the label is commonly applied to criminals who exhibit "predatory" behaviour, whether that behaviour results in death (e.g. serial killers), severe psychological harm (e.g. abusers, rapists, stalkers), or financial loss (fraudsters, burglars). Further, elaboration of the category has proceeded despite the obvious lack of formal similarity between human beings and e.g. polar bears or Venus flytraps (see Gentner and Asmuth 2019 on the metaphorical expansion of relational categories).

categorised as masculine nouns and are, therefore, expected to function as such through application of the appropriate endings.[36] Through this operation, the schema for masculine nouns of the kind Ice. nom.pl. -V_1*r* ~ acc.pl. -V_1 was elaborated in the following ways.

First, assignment of masculine grammatical gender to *iPad* and *bar* elaborated the schema for masculines so that it extended its remit to more nouns in nom.sg. -Ø. The lack of an ending in this form is relatively rare among masculine nouns and is mostly associated with stem-final *-l*, *-n*, *-r*, and *-s*, e.g. masculine *fugl* 'bird', *turn* 'tower', *vír* 'wire', *ís* 'ice' (see PII:215). Most masculine nouns end in either *-ur*, e.g. *hest**ur***, *gest**ur***, or *-i*, e.g. *grann**i*** 'neighbour', in the nominative singular.[37] Secondly, extension of the relational schema [XV_1*r*]$_{nom.pl.}$ ~ [XV_1]$_{acc.pl.}$ to include nom.pl. *iPadar* and *barir* served to elaborate the formal properties of the relevant schemas. In doing so, the emergent schemas are imbued with even more phonetic scope, further facilitating its extensibility. For this reason, the schema remains highly productive and is able to attract new items such as *iPad* and *bar*, facilitating their engagement in the relation type nom./acc.sg. -Ø ~ nom.pl. *-ar* ~ acc.pl. *-a*, cf. *iPad* ~ *iPad**ar*** ~ *iPad**a***, and nom./acc.sg. -Ø ~ nom.pl. *-ir* ~ acc.pl. *-i*, cf. *bar* ~ *bar**ir*** ~ *bar**i***. In turn, such developments further confirm and entrench the knowledge that endingless singular forms can alternate regularly with nom.pl. -V_1*r* and acc.pl. -V_1.

Thirdly, such instances of elaboration further entrench the relevant schema, in turn strengthening its representation in memory and, therefore, facilitating recall when the language user requires a point of reference for the inflection of a new masculine noun. Fourth, further entrenchment of the schema bolsters the cue validity that a form of the kind nom.pl. -V_1*r* has for another of the kind acc.pl. -V_1 — and vice versa. Further, due to the process of statistical learning, the more frequent extension of the alternation nom.pl. *-ar* ~ acc.pl. *-a* relative to that of nom.pl. *-ir* ~ acc.pl. *-i* is in large part a function of the higher type frequency of the former. Moreover, as the stems of nouns belonging to masculine classes in nom.pl. *-ar*, acc.pl. *-a* can be of practically any phonotactically permissible structure, their schemas are all but completely open. As noted in 3.2.3, high type frequency in conjunction with high schematicity can facilitate near limitless productivity.

[36] As far as I am aware, Ice. *bar* is the only masculine borrowing to have taken up the pattern nom.pl. *-ir* ~ acc.pl. *-i* (see also Kvaran 2005:345). Its assignment to the relevant class is probably due to phonetic similarity with e.g. *hv**al**ur* 'whale', *s**al**ur* 'hall', *d**al**ur*, 'valley, dale', *st**að**ur* 'place', *s**el**ur* 'seal', *r**ef**ur* 'fox', and others, all of which contain a rhyme in a lax vowel followed by a sonorant consonant and end in *-ir* in the nominative plural.

[37] See Þórhallsdóttir (1997) on evidence from language change for association of the ending nom.sg. *-ur* with masculine grammatical gender.

The arguments forwarded above support the core axiom of the usage-based cognitive approach reiterated in (14), above. Further, it identifies language use and, therefore, language change as functions of analogy. In other words, prior knowledge of relations and the argument structure that cause them to be perceived as similar facilitate extension of the relevant schematic knowledge to new contexts for the purpose of categorisation (see (17), above). Subsequent to extension, the schema is further elaborated to encompass new contexts of use, affecting the structure of schematic knowledge — linguistic or otherwise.

3.4 Chapter summary

The current chapter delineated the theoretical approach applied in Papers I–III in both general and context-specific terms. Section 3.1 compared the merits of and selected between opposing theoretical approaches to language change, i.e. generative and dual processing models, on the one hand, and the usage-based cognitive approach, on the other. In this connection, it was noted that the former posits genetically endowed language competence and assumes language change to be a function of grammar simplification, the mechanism for which is viewed as inter-generational transfer of the discerned grammar during first-language acquisition.

In order to identify the oversimplified view of language change as inherent to such claims, I presented examples of change that obviously complicates grammar, while further examples demonstrate that change can be a function of use among adult language users. It was also noted that while traditional generative and dual-processing approaches assume little role for the impact of frequency in language change, this position is incongruent with observable frequency effects. It was subsequently argued that the usage-based cognitive approach converges with the facts, e.g. that change occurs in adult language as well as in child language and that frequency impacts the process.

Section 3.2 fleshed out the tenets, principles, and axioms of the usage-based cognitive approach and simultaneously served as a brief state of the art. Terminology associated with the approach was defined in the context of Insular Nordic examples of morphological change. Subsection 3.2.1 argued in favour of rich memory for language through reference to different types of frequency effects in language use and change. Subsection 3.2.2 accounted for the impact of the causal relation *frequency* > *entrenchment* > *lexical strength* on the structure of grammar, before arguing that knowledge pertaining to frequency imbues memory representations for both specific and general experience. Subsection 3.2.3 examined the interaction of frequency with schematicity, where the latter serves to facilitate limited

productivity in the absence of significant type frequency. Subsection 3.2.4 discussed prototype structure as a property of categories, itself a function of cognitive economy. Subsection 3.2.5 offered a summary of Section 3.2. In Section 3.3, I discussed the assignment of linguistic and non-linguistic items to relational categories, the entrenchment of these in memory, and elaboration of relational schemas as a function of analogy.

4 Methodology

In this Chapter, I delineate the methodology employed for amassing and analysing the data presented in Papers I–III, in line with the theoretical tenets of the usage-based cognitive approach fleshed out in Chapter 3. Section 4.1 outlines the methodology employed for analysing the Icelandic data reported on in Paper I. Section 4.2 delineates the theoretical approach that determined the choice of methodological approach to the Icelandic data analysed in Paper II. Section 4.3 provides a detailed account of the methodological approach employed for analysing the data reported on in Paper III.

4.1 The methodology employed in Paper I

As reported in Paper I, according to counts based on the *IsTenTen* electronic corpus,[38] reanalysis of masculine plural *fætur* 'feet' as feminine occurs in the minority of instances. The gauge for the rate of reanalysis followed in Paper I is the number of individual occurrences of fem.pl.def. *fæturnar* (1,274 instances), masc.nom.pl.def. *fæturnir* (1,426), and masc.acc.pl.def. *fæturna* (2,882) (see PI:13–14). To determine the rate, the number of feminine forms is calculated as a percentage of the total number of instances (5,582 instances). Therefore, reanalysis occurs in 22.82% of instances.

In this connection, Paper I also identifies an interesting property of reanalysis pertaining to masculine forms in plural *-ur*: fem.pl.def. *fæturnar* is far more frequent than e.g. fem.pl.def. *veturnar* (of masculine *vetur* 'winter'). This fact is considered interesting because masc.acc.pl.def. *veturna* is much more frequent than both masc.nom.pl.def. *fæturnir* and masc.acc.pl.def. *fæturna* combined (see PI:14). At first glance this property of reanalysis might be considered understandable in light of the causal relation *frequency > entrenchment > lexical strength*. Indeed, according to the usage-based cognitive approach, masc.acc.pl.def. *veturna* is deterministically more entrenched on account of its relatively higher frequency and, therefore, logically has greater lexical strength than both masc.nom.pl.def. *fæturnir* and masc.acc.pl.def. *fæturna*. As a correlate of relative ease of access, we should expect the formal properties of frequent forms to be less subject to change than those of less frequent

[38] The *IsTenTen* corpus is an electronic text corpus for Icelandic. It is maintained by *Sketch Engine* and contains sources including social media. On my reasons for choosing to present results from this corpus rather than those returned from other electronic corpora for Icelandic, see Paper I (PI:23, endnote 18). See the bibliography for the relevant URL.

forms because ease of access implies less need to create a novel form in on-line language processing, which can happen when a given form is momentarily forgotten (see 3.2.1). All other things being equal, frequency counts from text corpora should either confirm or refute such correlation.

However, searches of the corpus for the individual forms listed above demonstrate a mismatch between the descending token frequencies of the respective feminine doublets relative to those of corresponding masculine forms for the same word. In other words, e.g. while masc.acc.pl.def. *veturna* is much more frequent than both masc.nom.pl.def. *fæturnir* and masc.acc.pl.def. *fæturna* combined, fem.nom./acc.pl.def. *fæturnar* is the most frequent feminine doublet. It is in light of this mismatch that schematicity, in addition to token frequency counts, is invoked as an explanatory device.

Paper I utilises Bybee's *Network Model* (e.g. 1985), with the employment of some innovative notational features. This method for modelling the emergent nature of form-function correspondences utilises connecting lines between segments or larger elements common to the words that form a given class. The lines are of varying thickness, depending on the nature of the correspondence — formal and/or functional (see PI:19; also 5.3.5). The greater the number of lines that connects the items in the network, the less schematic the set and the greater the likelihood that it is representative of an easily definable class.

Crucially, the greater the number of one-to-one correspondences is shown in Paper I to correlate with the impact of minimal schematicity as a facilitator of limited productivity. Through employment of the model and the innovative features in Paper I, it proves possible to not only demarcate the emergent status of specific sequences as representative of grammatical function, cf. the association of the sequence *-ur* with the function *nominative/accusative plural*, but also to model the degree to which other sequences within a word align with prototypical formal attributes of a microclass within a subset of strong feminine nouns in plural *-ur*. The number of connections denoting a form-function match between individual forms in the network is argued to correlate with the potential for alignment of those forms with a specific schema (see the network in PI:19 and those represented in 5.3.5).

Thus, the network presented in Paper I shows that plural *fætur* is the only original masculine form in *-ur* that aligns perfectly with the subordinate level schema $[X\text{æ}T\textit{ur}]_{\text{nom./acc.pl.}}$, which was posited for functionally equivalent forms of the feminine Xó/æT-microclass, i.e. Ice. plural *blækur* 'nonentities, wretches', *bækur* 'books', *bætur* 'patches, remedies, compensation, financial aid', *brækur* 'trousers', *nætur* '(fishing) nets',

rætur ‘roots’. Additionally, plural *fætur* also aligns with the basic level schema [X*ur*]$_{\text{nom./acc.pl.}}$ on account of the ending plural *-ur* alone (see 3.2.4 on taxonomic relations between the constituent schemas of categories; see 5.2 and 5.3.2 on the prototype structure and taxonomic structure of the subtype for Icelandic feminine nouns in plural *-ur* specifically).

Conversely, plural *vetur* aligns at best only tentatively with subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ (see PI:19–20). Therefore, the network demonstrates that plural *vetur* is at best formally ambiguous with regard to class membership, even though it is easily associated with feminine grammatical gender due to alignment with basic level [X*ur*]$_{\text{nom./acc.pl.}}$.[39] Thus, plural *fætur* can be specifically attracted by two schemas, each representative of its own level of abstraction. Therefore, it is argued that frequency counts from the *IsTenTen* corpus provide support for the usage-based cognitive approach for the following reasons.

First, in line with the posited impact that frequency has on entrenchment and lexical strength, we should expect the attractive force of a schema that abstracts over very few items to be limited — all other things being equal. Accordingly, as noted above, plural *fætur* is reanalysed as feminine in the minority of cases. However, this then begs the following questions: Why is fem.pl.def. *fæturnar* markedly more frequent than any functionally equivalent feminine doublet of an original masculine form in plural *-ur*? Given that any form with the ending in question can be aligned with medially frequent [X*ur*]$_{\text{nom./acc.pl.}}$, do the rates of reanalysis reported in Paper I not suggest that low frequency [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ nonetheless exhibits greater productivity? It is in this context that the condition “all other things being equal” proves highly relevant.

The answer to the questions just posed brings us to our second reason, namely that low frequency can interact with low schematicity to facilitate the gang effect as a function of prototype structure (see 3.2.3 and 3.2.4). Thus, employment of the network model in Paper I demonstrates that alignment with subordinate level, low frequency [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ facilitates assignment of plural *fætur* to the X*ó*/*æ*T-microclass and, therefore, identity with specifically feminine formal attributes beyond plural *-ur*. For this reason, plural *fætur* is discernibly far less ambiguously feminine in form than plural *vetur*.

As is argued in Papers I and II, frequency counts also shed light on the nature of reanalysis within the context of the paradigm as a whole. Both papers account for reanalysis as extension, by which formal and/or functional similarity across intra-paradigmatic relations is aligned as a function of parallel connectivity. Thus, it is argued, the cue validity of plural

[39] See 5.3.2 on the status of the sequence *-ur* in forms such as *bræð**ur*** ‘brothers’, *fing**ur*** ‘fingers’, and *vet**ur***.

fætur suffices to motivate deduction of innovative fem.pl.def. *fæturnar*, in line with existing relations such as feminine plural *rætur* ~ fem.pl.def. *ræturnar*. However, frequency counts for other forms of Ice. *fótur* reveal that singular forms of the word hardly — if ever — occur in an overtly feminine form (PI:8). Therefore, this property of usage is considered evidence for reanalysis as a two-step process, by which individual forms have cue validity for the treatment of certain others within the paradigm, as opposed to a wholesale process that affects all forms simultaneously.

Following from the above discussion, a methodological approach that employs frequency as diagnostic of productivity supports rich memory for language. As argued in Paper I, it is likely overtly masculine attributes exhibited by singular forms of the noun Ice. *fótur* that impede reanalysis as feminine. By the same token, plural *fætur* patterns perfectly with other feminine forms at two distinct levels of abstraction (see above). Therefore, the formal attributes in question, as well as their dispersion over gender-determined contexts, must have representation in memory — otherwise, we should not expect to find a correlation between attributes' force in numbers and the rate of reanalysis.[40]

4.2 The methodology employed in Paper II

As noted in Chapter 1, Papers I and II are related in terms of thematic content: Both deal with gender reanalysis and assignment of grammatical gender on account of similarity to feminine nouns in plural *-ur*. However, unlike Paper I, Paper II seeks to discern the different analogical means by which different forms of the Icelandic borrowings *blók* 'wretch, nonentity' and *kók* 'Coke™' are deduced, i.e. by intra-paradigmatic levelling and extension. The different mechanisms of change are accounted for via reference to non-linguistic factors such as frequency and schematicity, on the one hand, and semantics and pragmatics, on the other.

The methodology employed in Paper II seeks to highlight the basic forms for levelling in terms of a **relation between meaning and form**, which, as demonstrated by Bybee (1985), is reflected cross-linguistically in paradigmatic structure. Synchronically, the relation between meaning and form is betrayed by formal coherence among forms that share a semantically highly significant morphological value, such as those defined in terms of tense and aspect among verbs, in terms of number among nouns. By the same token, the tendency towards phonetic *coherence* as a correlation of shared semantics can lead to formal *distinction* between forms on different sides of the significant semantic divide within the

[40] In this connection, see 3.2.1 on the development of Eng. *keep, mean, sleep* vs. that of *creep, leap, weep*.

paradigm. Diachronically, the relation between meaning and form is manifested through levelling as forms that share a significant semantic property are levelled to resemble each other more than those forms that express an opposing value (see e.g. Lahiri 2000; also PIII).[41]

Informed by this approach to meaning and form, Paper II argues that innovative forms of Ice. *kók*, which usually shows agreement as neuter in the mass noun sense, are deduced via different analogical means, i.e. extension or levelling. Further, application of either mechanism depends on the degree of semantic relatedness to the source of innovation, i.e. nom./acc.sg. *kók*. On the basis of numerical evidence, Paper II demonstrates that the formal distinction between nom./acc.sg. *kók* and innovative feminine plural *kækur*, when the latter occurs instead of plural *kók*, aligns with the semantic difference between singular and plural, as mirrored in functionally equivalent relations of the X*ó*/*æ*T-microclass, cf. e.g. *bók* ~ *bækur*, *rót* ~ *rætur*, etc.

Conversely, levelling to fem. dat.sg. *kók*, instead of neut. dat.sg. *kóki*, where the latter is most common in the mass noun sense, reflects the semantic relatedness of the relevant forms one to another, cf. the source nom./acc.sg. *kók* as used in the count noun sense. Crucially, syncretism in nominative/accusative/dative singular is a highly frequent attribute of strong feminine nouns in Icelandic, cf. e.g. nom./acc./dat.pl. fem. *bók* 'book', *mynd* 'picture, form', *skeið* 'spoon', suggesting the influence of type frequency and dispersion on the direction of levelling. If all members of the paradigm for Ice. *kók* were reanalysed as feminine simultaneously, we might expect feminine forms to occur at proportionately representative rates compared with neuter doublets. However, this is not the case. Paper II therefore relies on frequency counts from text corpora in order to shed light on the nature of reanalysis as a function of distinct analogical means, both of which presume influence from respective forms of the word (see PII:206).

Paper II also seeks to demonstrate that innovative (non-humorous) plural *blækur* and (humorous) plural *kækur* are deduced from singular *blók* and *kók*, respectively, by identical analogical means. This endeavour is a reaction to the belief expressed by some that plural *kækur* is a joke and, therefore, not an example of "real" language use, while plural *blækur* is considered to reflect "real" usage (see an example of this view in PII:195–196, footnote 1). In

[41] This relation between synchrony and diachrony exemplifies Bybee's view that domain-general cognitive mechanisms of language change give rise to "universal" tendencies through language use, although specific outcomes of the process are not universal as they are constrained by the respective grammars of individual languages (Bybee 2008:120–121).

order to demonstrate that the cognitive mechanisms for deduction of the relevant forms are not substantively different, Paper II demonstrates that analogy feeds on both semantics and pragmatics to extend knowledge of contextually appropriate language use to imagined, inappropriate contexts. The result of extension to this new context can seem so absurd as to imbue to the use of language in it humorous. Paper II concludes that, no matter the motivation for extension, language use in new contexts is the result of analogy and that this is the cognitive means by which all innovative forms of Ice. *blók* and *kók* are deduced. Therefore, if use of plural *blækur* represents "real" language use, so does use of e.g. plural *kækur*.

4.3 The methodology employed in Paper III

Evidence from language change suggests time and again that the basic forms of paradigms are also those with the highest token frequency — all other things being equal. In this connection, Paper III utilises frequency counts from an electronic text corpus for Modern Faroese, *Teldutøka tekstasavn Føroyamálsdeildarinnar* (TTF),[42] to demonstrate that this property of usage facilitates deduction of innovative dat.pl. *firðum*, which is based on by far the most frequent form of Far. *fjørður*, i.e. dat.sg. ***firð****i*. However, while original dat.sg. *firði* has likely always been the most frequent form (see below), innovative forms containing the variant *fjørð-*, cf. innovative dat.sg. *fjørði*, are based on far less frequent forms, such as nom.sg. ***fjørð****ur*, acc.sg. ***fjørð*** and/or original dat.pl. ***fjørð****um*.

At first glance, the point just made might appear to contradict arguments for the impact of the causal relation *frequency > entrenchment > lexical strength* as a determinant of the direction of levelling. In other words: Why should a stem variant found in a far less frequent form of the paradigm be selected at the cost of a variant found in by far the most frequent form, given that the latter has the greatest lexical strength? In answer to this seeming contradiction, Paper III argues for the impact of two dimensions of frequency on the direction of levelling. First, token frequency is demonstrably deterministic of lexical strength and serves as the catalyst of levelling, given a specific context of use. Crucially, token frequency is blind to intra-paradigmatic dynamics between forms with regard to the relation between meaning and form (see below). Secondly, in the wake of change predicated on token frequency, it is argued that language users relied on the dispersion of stem variants across the

[42] See the bibliography for the relevant URL.

singular and plural portions of the paradigm in order to establish a relation between either value and a respective stem variant.

In context-specific terms, Paper III argues that the stem variant *firð-* is basic to the paradigm, due to its high token frequency and the impact of this factor on lexical strength. Further, it is hypothesised that the same form has likely always been basic to Far. *fjørður*, as its number of referents and their individual meanings have not changed during the centuries since settlement of the Faroe Islands. Therefore, it is concluded that different rates of usage for individual forms of the word have always correlated proportionately with those reflected by corpus data for the modern language, as usage-needs have not changed substantially with regard to the number of definitions.

Subsequently, extension of the variant *fjørð-* was not a function of token frequency, but rather constituted a reactionary attempt to establish opposing relations between meaning and form across the semantic divide singular vs. plural. In other words, the plural was represented by *firð-*, cf. nom./acc.pl. ***firð****ir*, innovative dat.pl. ***firðum*** (beside original *fjørðum*) and singular by the variant *fjørð-*, cf. nom.sg. ***fjørð****ur*, acc.sg. ***fjørð***, innovative dat.sg. ***fjørð****i* (beside original *firði*). It is the often-observed course of levelling, which time and again proceeds from the most frequent member, that permits positing this chronology. Thus, frequency counts based on text corpora guide the application of usage-based theory to the available data, rather than the other way around.[43]

In the absence of written sources for different stages of Faroese language history (see discussion in Chapter 2), Paper III also employs a method of determining historical token frequencies of different forms of the noun *vøllur*. The reason for this endeavour is that, unlike *fjørður*, Far. *vøllur* has innovated with regard to its number of referents since the time of settlement. Therefore, given the usage-based cognitive assumption that frequency is a determinant of the direction of levelling, it proved necessary to ascertain which form(s) most likely served as basic before the process commenced. To this end, Paper III applies a comparison of frequency data obtained from *TTF* and *Íslenskt textasafn* (*ÍT*),[44] an electronic text corpus for Icelandic, spanning all periods.

The utility of *ÍT* in estimating historical frequencies for earlier Faroese is obviated by the discussion of Old West Nordic in Chapter 2. In other words, Icelandic and Faroese were

[43] Paper III also relies on frequency counts from the same corpus to determine the direction of levelling taken in the paradigm of Far. *vøllur*. The methodology for achieving that endeavour is treated specifically in section 4.3.1.

[44] See references for electronic access to the corpus.

almost indistinguishable in the centuries immediately subsequent to settlement, so use of the search function *Fornrit* ‘ancient texts’ likely gives an indication of usage patterns for search words in Old West Nordic. Crucially, however, the usage-based cognitive approach acknowledges that language users’ perception of the real-world conditions that prevail in their environment impact usage. Thus, due to both weather and topographical traits of the Faroe Islands, pre-20th century referents of Far. *vøllur* — before its use in the sense ‘(sports) pitch’ increased and the sense ‘airport’ came into use — tended to be small and, therefore, not to perform established functions. Conversely, referents of the Icelandic cognate *völlur* are typically larger, likelier to perform established functions and, therefore, occur more often.

Based on this comparison, the estimation of historical token frequencies for individual forms of Far. *vøllur* asserted that the word occurred less frequently in the sense ‘(sports) pitch’ in Faroese than in Icelandic. However, I decided to be generous with the estimate and assumed half the token frequency of the Icelandic equivalent for pre-20th century Faroese. Additionally, the number of instances in which Far. *vøllur* occurred in the sense ‘airport’ according to *TTF* were subtracted for earlier phases of the language. Use of the word in the senses ‘field, grassy ledge on a rock face’ was considered to proportionally represent historical token frequency, given that the number of referents in these senses have likely changed little over the centuries. On this basis, it is estimated that acc.sg. *vøll* was the most frequent form of the word in pre-20th century Faroese (see PIII:68, Table 4). Given the relation *frequency* > *entrenchment* > *lexical strength*, Paper III assumes that this conclusion will have had consequences for the choice of basic form of Far. *vøllur* and determined the direction of levelling within the paradigm.

5 Icelandic data: Papers I and II

This chapter details the respective studies presented in Papers I and II, both of which account for the productivity of the X*ó*/*æ*T-microclass — albeit to varying degrees of focus. Due to the common thematic content, both studies can mostly be delineated within the same sections but are also treated separately where necessary. Section 5.1 fleshes out the contents of Papers I and II in light of theoretical focus. In Section 5.2, I delineate the prototype structure that centres on the feminine X*ó*/*æ*T-microclass.

Section 5.3 fleshes out the analysis presented in Paper I, accounting for the linguistic and non-linguistic factors that facilitate the reanalysis of masculine forms in plural *-ur* as feminine at different rates of frequency. To this end, subsection 5.3.1 demonstrates that frequency of use impacts categorisation via the domain-general cognitive process of statistical learning, the theoretical basis for which was discussed in 3.2.2. In subsection 5.3.2, I elaborate on cognitive economy as a property of functionally related linguistic categories. Subsection 5.3.3 presents instances of actual language use as evidence for hierarchical relations between varyingly schematic representations of linguistic categories. In subsection 5.3.4, I discuss the effects of minimal schematicity as a facilitator of limited productivity in light of the arguments forwarded in the preceding subsections. Subsection 5.3.5 models the extent to which varying degrees of schematicity facilitate different rates of productivity for functionally equivalent schemas represented at different levels of abstraction, where the notational conventions employed characterise what I call the 'net effect'. Subsection 5.3.6 offers a summary of Section 5.3.

Section 5.4 fleshes out the analysis presented in Paper II. Subsection 5.4.1 presents the content of Paper II. In subsection 5.4.2, I argue for the deduction of plural *kækur* and *blækur* from singular forms of the respective paradigms as part a two-step process. Subsection 5.4.3 details the different analogical means by which individual forms of a paradigm can be deduced, i.e. by extension and levelling. Subsection 5.4.4 argues that plural (non-humorous) *blækur* and (humorous) *kækur* are deduced by extension, negating the position that one is a better example of "real" language use than the other. Section 5.4.5 provides a summary of Section 5.4. Section 5.5 offers a summary of the current chapter.

5.1 Content and focus of Papers I and II

Paper I accounts for the different rates at which Icelandic masculine forms in plural *-ur* are reanalysed as grammatically feminine on the basis of schematicity. Specifically, it establishes a correlation between the rate of reanalysis, the dispersion of plural *-ur*, and the degree of formal similarity that individual masculine forms exhibit to a phonetically coherent prototype in the same ending. The prototype is manifested by the individual members of the X*ó*/*æ*T-microclass (following Dressler 2003), i.e. the feminine nouns Ice. *blók* 'non-entity, wretch', *bók* 'book', *bót* 'patch, remedy', *brók* 'trousers', *nót* '(fishing) net'. *rót* 'root'. While the X*ó*/*æ*T-microclass is considered minimally schematic, the masculine forms discussed in Paper I are as phonetically dissimilar beyond plural *-ur* as e.g. masc. plural *bræður*, *eigendur*, *fingur*, *fætur*, *vetur*.

The occurrence of reanalysis can be discerned when masculine forms in plural *-ur* acquire doublets containing the definite article fem. pl. *-nar*. Examples are *bræður**nar***, *eigendur**nar***, *fingur**nar***, *fætur**nar***, *vetur**nar***, which exist alongside (original/standard) masc. nom.pl.**def**. *bræður**nir*** ~ acc.pl.**def**. *bræður**na***, *eigendur**nir***/***-na***, *fingur**nir***/***-na***, *fætur**nir***/***-na***, *vetur**nir***/***-na***. Crucially, the feminine doublets occur in the minority of cases (see PI:10–11). As explained in Chapter 4, in order to determine the rate of reanalysis, Paper I adds the token frequencies of the masculine and the feminine forms together before the number of feminines is calculated as a percentage of that total.

Paper I argues that reanalysis occurs — all other things being equal — due to the dispersion of plural *-ur*: despite the ending occurring in 14.92% of noun paradigms, 91.89% of these are feminine. However, the analysis demonstrates that all things are *not* equal, highlighting different rates of reanalysis as a function of the degree to which a masculine noun in plural *-ur* aligns with the functionally equivalent schema for the X*ó*/*æ*T-microclass. In other words: the higher the number of one-to-one formal and functional correspondences between forms, the greater the likelihood that a masculine form in plural *-ur* will be reanalysed as feminine on alignment with the minimally schematic schema $[X\textit{æ}T\textit{ur}]_{\text{nom./acc.pl.}}$. Therefore, as discussed in 3.2.3, minimal schematicity correlates with a boost to — albeit still limited — productivity as a function of the gang effect.

Paper II deals with the limited productivity of the X*ó*/*æ*T-microclass specifically. Despite low type frequency and low schematicity, the relevant schemas have been extended to three new contexts since the 16$^{\text{th}}$ century, i.e. innovative feminine pl.def. *fæturnar*; feminine dat.sg. *kók*, gen.sg. *kókar*, nom./acc.pl. *kækur*, cf. the borrowing *kók* 'Coke™', which is mainly neuter in the mass noun sense but feminine in the count noun sense; and the

entire paradigm of borrowed feminine *blók*. The three paradigms are shown in (18), beside that of *rót* (for comparison, the nominative and accusative plural forms of the **article** are included for the relevant forms of feminine *rót* and masculine *fótur*).

(18)			fem.	fem.	neut./fem.[45]	masc./fem.
	sg.	nom.	*rót*	*blók*	*kók*	*fótur*
		acc.	*rót*	*blók*	*kók*	*fót*
		dat.	*rót*	*blók*	*kóki*/*kók*	*fæti*
		gen.	*rótar*	*blókar*	*kóks*/*kókar*	*fótar*
	pl.	nom.(**def**.)	*ræt**ur***	*blækur*/*blókir*	*kók*/*kækur*	*fæt**urnir***/***-nar***
		acc.(**def**.)	*ræt**ur***	*blækur*/*blókir*	*kók*/*kækur*	*fæt**urna***/***-nar***
		dat.	*rótum*	*blókum*	*kókum*	*fótum*
		gen.	*róta*	*blóka*	*kóka*	*fóta*

Papers I and II both demonstrate that extension of the X*ó*/*æ*T-microclass pattern to new contexts is a function of phonetic and semantic coherence between the sources of innovation and the respective feminine targets listed above, cf. plural *bræð**ur**nar*, *eigend**ur**nar*, *fingur**nar***, *fæt**ur**nar*, *vet**ur**nar*** (Paper I); plural *kækur*, *blækur* (Paper II). However, before writing of Paper II commenced, I had become aware through informal conversations with colleagues that some do not view plural *kækur* as an example of "real language use". Conversely, the same did not apply to plural *blækur*. This position appears to be based on the fact that *kækur* is used in humorous contexts only (see PII:195, footnote 1). Therefore, Paper II seeks to demonstrate through joint appeal to semantics, pragmatics, and schematicity that deduction of plural *kækur* on the basis of sg. *kók* by no means belies "real language use". Rather, it concludes that all linguistic innovation predicated on analogy is a property of language use.

In order to account for the feminine innovations discussed, appeal is made in Papers I and II to the domain-general cognitive processes of statistical learning and categorisation, as

[45] Although plural *kók* is almost always qualified by feminine forms, some examples of neuter modifiers can be found, cf. (**neut**.**acc**.) ***tvö*** *kók* 'two Cokes™'. Further, the form pl. *kók* patterns with most other strong neuters in that it is fully syncretic with the nominative/accusative singular, i.e. *kók*, e.g. neut. nom./acc.sg., nom./acc.pl. *borð* 'table(s)', *hús* 'house(s)', *epli* 'apple(s)'. This suggests that plural *kók* is phonetically based on the singular mass noun sense of the word, despite being used as an uninflected feminine form in most instances in the plural count noun sense (see PII:207).

guided by analogy (see 3.2.2 and 3.3). In Paper I, such appeal proves a means to demonstrate that a taxonomy of formally distinct, yet functionally continuous schemas has been elaborated via categorisation and entrenchment. Further, both studies argue that categorisation by the minimally abstract sister schemas $[X\acute{o}T]_{\text{nom./acc./dat.sg.}}$ ~ $[X\text{æ}T\textit{ur}]_{\text{nom./acc.pl.}}$ ~ $[X\text{æ}T\textit{urnar}]_{\text{nom./acc.pl.def.}}$ facilitates perception of one-to-one formal correspondences between inflectional forms, yielding parallel connectivity between functionally equivalent arguments on each side of a relation (see discussion and examples of parallel connectivity in Section 3.3). Papers I and II argue that this process results in alignment with relations represented by the sister schemas above, ultimately motivating assignment of the relevant forms to the X*ó*/*æ*T-microclass specifically.

Further, Papers I and II demonstrate that the greater the extent of one-to-one formal and functional correspondences between aligned forms, the greater the likelihood that an existing form be assigned feminine grammatical gender. In Paper I, the greater likelihood of reanalysis as feminine is reflected by the mismatched descending frequencies between functionally equivalent sets of masculine forms, on the one hand, and innovative, feminine doublets, on the other (see above; also PI:14, Table 1). Ultimately, a greater degree of one-to-one correspondence accounts for both the limited productivity of the feminine subtype in plural *-ur* and the higher rate at which plural *fætur* is reanalysed as feminine compared with other masculines in the same ending.

In Paper II, recourse is made to both extension and levelling in accounting for the deduction of feminine forms of borrowed *blók* and *kók*. Further, it is demonstrated that semantic relationships between specific forms within each paradigm, real world referents of each word, and use in context, i.e. pragmatics, motivate the analogical means of deduction. In terms of semantics, for example, Ice. *blók* can refer to people of any gender, a property of usage that likely facilitates application of feminine grammatical gender.[46] In terms of pragmatics, Ice. *kók* often occurs in a syntactic context where it is qualified by feminine modifiers, used to qualify an omitted feminine noun that names the container or vessel in which the drink is bought or served (see PII:202). It is argued that the distinct semantic and

[46] In this connection, Icelandic does exhibit a tendency to align biological sex and grammatical gender. Consider examples such as *bróðir* ‘brother’ and *móðir* ‘mother’. The two words inflect identically in their indefinite forms, a remnant from the Proto-Indo-European gender system that distinguished animate nouns from inanimate (this is a gross simplification of the situation). However, despite following identical patterns of inflection, *bróðir* is grammatically masculine while *móðir* is grammatically feminine. For an elaboration on the link between biological sex and grammatical gender in Icelandic (see PII:214–215).

pragmatic factors at play, in conjunction with schematicity, facilitate graded membership of the X*ó*/*æ*T-microclass.

In light of the above, the common themes of Papers I and II are the following:

1. Both papers detail innovation and change pertaining to the gender assignment of Icelandic nouns as a function of analogical reasoning.
2. Both papers demonstrate the influence of minimal schematicity as a facilitator of limited productivity.
3. In both papers, reanalysis as feminine is attributed to the high rate at which plural *-ur* is dispersed among classes of feminine nouns — all other things being equal.
4. In both papers, semantic and pragmatic properties of the forms under analysis are shown to impact both the rate of alignment with the X*ó*/*æ*T-microclass and, consequently, the rate of reanalysis as feminine.
5. In both papers, frequency of use is a gauge for the influence of minimal schematicity as a facilitator of productivity.

The rest of the current chapter elaborates on the arguments forwarded, the theoretical approach applied, the methodology employed, the data presented, and the conclusions made in Papers I and II. The next section delineates the prototype structure of the feminine subtype in plural *-ur* as it centres around the X*ó*/*æ*T-microclass.

5.2 Prototype structure that centres on the Xó/æT-microclass

In subsection 3.2.4, I applied Rosch's principle of cognitive economy to delineate a taxonomy of functionally related schemas that represent members of the category FURNITURE at different levels of abstraction, cf. the superordinate category label FURNITURE, basic CHAIR, and subordinate DINING CHAIR. In the current section, I delineate the prototype structure of the Icelandic feminine subtype in plural *-ur* according to the same principle, accounting for its convergence with the members of the X*ó*/*æ*T-microclass at the subordinate level. At both the basic and the subordinate levels, the taxonomy is defined by varyingly schematic sets of feminines in plural *-ur*, represented by the schema [X***ur***]$_{\text{nom./acc.pl.}}$, cf. e.g.

plural *bæk**ur***, *geit**ur***, *stelp**ur***, and [X***æ***T***ur***]$_{\text{nom./acc.pl.}}$, cf. e.g. *bæk**ur***, *næt**ur***, *ræt**ur***, respectively.[47]

Like the majority of Icelandic strong feminine nouns (see footnote 4), members of the subtype in plural *-ur* comprise a single syllable in the syncretic nominative/accusative/dative singular, which alternates in almost all cases with a disyllabic genitive singular in *-ar* and a syncretic, disyllabic nominative/accusative plural in *-ur*, e.g. nom./acc./dat.(/gen.)sg. *kind*(*ar*) 'sheep' ~ plural *kindur*, *geit*(*ar*) 'goat' ~ *geitur*, *eik*(*ar*) 'oak' ~ *eikur*.[48] Further, these alternate intra-paradigmatically with a definite form in feminine *-nar*, cf. sg. *kind*(*ar*) ~ pl.def. *kindurnar*, *geit*(*ar*) ~ *geiturnar*, *eik*(*ar*/*-ur*) ~ *eikurnar*. Additionally, many paradigms exhibit vocalic alternation stemming from Proto-Nordic *i*-umlaut, by which stem vowels were fronted due to the influence of unstressed **i* (which has since been lost from the following syllable): nom./acc./dat.sg. *mörk* '250 gr.' ~ gen.sg. *m*a̲*rkar*/*m*e̲*rkur* ~ plural *m*e̲*rkur*, where PNc. **a* was original, sg. *kló*(*ar*) 'claw' ~ plural *kl*œ̲*r*,[49] *r*ó̲*t*(*ar*) ~ *r*œ̲*tur*, where PNc. **ō* was original.

The feminine subtype in plural *-ur* can be demarcated according to a continuum of prototypicality, determined in part by the morphophonological attributes of the paradigms to which the relevant schemas have been extended as graded similarity exhibits family resemblance structure (see 3.3; also 5.3.1). In this connection, Papers I and II detail the extension of the relevant schemas to the paradigm of masculine *fótur*, cf. fem.pl.def. *fæturnar*, and those of borrowed *blók*, cf. pl. *blækur*, and *kók*, cf. pl. *kækur*. Thus, it is argued that the nominative/accusative/dative singular forms of a prototypical subset of feminines within the subtype each comprises a single syllable with the diphthong *ó* [ou:] as its nucleus,[50] and a coda in either *t* or *k*.[51] Further, *-ar* is the prototypical ending of the genitive singular form, as is the case for the vast majority of Icelandic strong feminines.

---

[47] In the schema [X*ur*]$_{\text{nom./acc.pl.}}$, 'X' stands for 'a stem of any shape preceding the ending plural *-ur*'. In the schema [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, 'X' generalises over the stem-initial consonant (cluster) of members of the X*ó*/*æ*T-microclass, i.e. *b*, *br*, *n*, and *r*, cf. ***b****ækur*, ***b****ætur*, ***br****ækur*, ***n****ætur*, ***r****ætur*, (possibly *d*, cf. ***d****ætur* 'daughters', and *m*, cf. ***m****æður*; see below), before the addition of plural ***bl****ækur* and ***k****ækur*. Further, the notation 'T' represents the consonants *k* and *t*, cf. *bæ****k****ur*, *bæ****t****ur*, *bræ****k****ur*, *næ****t****ur*, *ræ****t****ur*, *blæ****k****ur*, and *kæ****k****ur* (possibly also *ð*; see footnote 51). Thus, 'X*æ*T' stands for 'stems like *bæk-* and *ræt-* preceding the ending plural *-ur*'.

[48] The variant gen.sg. *eikur* exists beside gen.sg. *eikar*.

[49] The ending *-r* of plural *klæ****r*** is the result of a historical process of contraction from earlier **-ur*.

[50] The general rule for vowel length in Icelandic is that stressed vowels are long in open syllables, but short in closed syllables. Stress always falls on the initial syllable (see Árnason 2005:135).

[51] After publication of Papers I and II, it was brough to my attention that the feminine noun *glóð* 'ember' alternates with plural *glæður*, as well as plural *glóðir* (I had only heard and only ever used the latter). Clearly,

Moreover, as is the case for the subtype in plural *-ur* where applicable, *i*-umlaut alternation is applied in the inflection of those nouns that constitute the prototype, cf. nom./acc./dat./gen.sg., dat./gen.pl. *-ó-* [ou:] and nom./acc.pl. *-æ-* [ai:], e.g. *bók- ~ bækur*, *rót- ~ rætur*. This morphophonological attribute is also associated with the extension of the subtype's schema(s) to new contexts, cf. nom./acc./dat./gen.sg., dat./gen.pl. *blók- ~* plural *blækur* (see PI:3–5). Further, although the innovative feminine form pl.def. *fæturnar* does not involve *extension* of the alternation *ó ~ æ* — the [ai:] of plural *fætur* is the result of historical fronting of PNc. **ō* by *i*-umlaut — alternation between *ó* and *æ* within the paradigm of masculine *fótur* is likely a motivating factor in the relatively high rate at which plural *fætur* is attracted by the X*ó*/*æ*T-microclass.

Given this morphophonological definition imposed on the prototype, feminine (borrowed) *blók*, (native) *bók*, *bót*, *brók*, *nót*, and *rót*, cf. sg. *blók*(*ar*) ~ nom./acc.pl. (standard) *blækur*, *bók*(*ar*) ~ *bækur*, *bót*(*ar*) ~ *bætur*, *bró*k(*ar*) ~ *brækur*, *nót*(*ar*) ~ *nætur*, *rót*(*ar*) ~ *rætur*, are considered prototypical of the feminine subtype in plural *-ur* and a model for the analogical extension of its schemas. Further, compounds with Ice. *-rót* as their final constituent, such as *engiferrót* 'ginger', i.e. *engifer-rót*, *kvaðratrót* 'square root', *gulrót* 'carrot', *piparrót* 'horseradish', also exist. Despite the obvious association of their latter constituent with *rót*, such compounds are here considered distinct lexical items that boost the type frequency of the feminine subtype in plural *-ur*. This is due to the fact that the referent of each compound renders it lexically distinct, both from *rót* and each other. In other words, one cannot refer to a carrot or ginger as *rót* alone and expect language users to discern the referent.

By the same token, compounds in *-bók*, such as *kirkjubók* 'church book', i.e. *kirkju-bók*, *dagbók* 'diary', *fundarbók* 'book of minutes', do *not* boost the type frequency of the subtype: all are types of *bók* that can be referred to as such without further specification within the appropriate context. For this reason, however, it must be conceded that compounds in *-bók* contribute to the *token* frequency of Ice. *bók*. The same applies to compounds in plural *-bætur*, cf. *atvinnuleysisbætur* 'unemployment benefit', i.e. *atvinnuleysis-bætur*, *húsnæðisbætur* 'rent rebate', i.e. *húsnæðis-bætur*, which are kinds of financial aid; in *-brók*,

alternation of the kind sg. *glóð* ~ pl. *glæður* accords with the pattern of alternation exhibited by e.g. *bók ~ bækur*, *rót ~ rætur* and, therefore, elaborates the relevant schemas to include one example of stem-final *ð* (potentially also plural *mæður*; see below). Elaboration of the schema to include stem-final *ð* presents no giant feat for generalisation and does not create problems for any of the claims made in Papers I and II.

cf. *nábrók* ‘necro-pants’, i.e. *ná-brók*, which are a kind of trousers; *-blók*, cf. e.g. *skrifstofublók* ‘pencil-pusher’, i.e. *skrifstofu-blók*, which is a kind of non-entity.[52]

A less central member of the subtype relative to the definition of its prototype is feminine *nótt* ‘night’. The vowel of the singular is short due to the nature of its coda (see footnote 50): Icelandic orthographic <tt> represents [ht],[53] cf. nom./acc./dat.sg. *nótt* [nouht]. Further, the *nótt* paradigm also contains the form gen.sg. *nætur*, which is syncretic with the plural form, i.e. plural *nætur* (both of which are, in turn, syncretic with the plural of Ice. *nót* ‘(fishing) net’). This is a relatively rare iteration of syncretism. Despite these non-prototypical attributes, Ice. *nótt* is considered to straddle the cusp of prototypicality as a satellite, due to the nature of vocalic alternation between *ó* and *æ* manifested by its paradigm, in conjunction with the ending plural *-ur*. For the purpose of comparison, the full paradigms for the X*ó*/*æ*T-microclass members *rót* and *bók*, as well as near-prototypical *nótt*, are shown in (19).

(19)

sg.	nom.	*rót*	*bók*	*nótt*
	acc.	*rót*	*bók*	*nótt*
	dat.	*rót*	*bók*	*nótt*
	gen.	*rótar*	*bókar*	*nætur*
pl.	nom./acc.	*rætur*	*bækur*	*nætur*
	dat.	*rótum*	*bókum*	*nóttum*
	gen.	*róta*	*bóka*	*nótta*

Still more distant satellites are feminine *geit* and *eik*. Although the coda of each complies with the phonological definition of the prototype, the nucleus of both words consists of the diphthong [ei:]. This vowel does not engage in *i*-umlaut alternation, cf. sg. *g**ei**t*(*ar*) ~ nom./acc.pl. *g**ei**tur*, as opposed to e.g. prototypical *bók*(*ar*) ~ *bækur*. Further, the existence of the doublet form gen.sg. *eikur*, beside *eikar*, further distances the morphophonological attributes of *eik* from those of the prototype. The same can be said of nouns such as *brík* ‘armrest, bracket’, *flík* ‘item of clothing’, and *tík* ‘female dog, (derogative) bitch’, which have long had doublets in the genitive singular, cf. *bríkur*/*bríkar*, *flíkur*/*flíkar*, *tíkur*/*tíkar* (see

[52] Conversely, as compounds in *-blók*, *-bók*, *-bót*, *-brók*, and *-rót* do not comprise a single syllable, they are not considered to boost the type frequency of the prototype.

[53] See Árnason 2005:161, 206–207 on phonetic transcription conventions for Icelandic consonants.

Guðmundsson 1922:69), as is also true of *flík* in the nominative/accusative plural, cf. *flíkur*/*flíkar* (see above about gen.sg., plural *nætur*).[54] The feminine noun Ice. *vík* has gen.sg. *-ur* only (see Jónsdóttir 2020:24 and sources cited there).

Less peripheral on formal grounds are plural *dætur* (of *dóttir* 'daughter') and relatively more peripheral plural *mæður* (of *móðir* 'mother'). While *dætur* meets the phonological definition of the Xó/æT-microclass in terms of its rhyme, *mæður* is considered more peripheral on account of the onset of its second syllable, i.e. inter-vocalic <ð> [ð] (though, see footnote 51). Further, it is likely that *dóttir* and *móðir* form a more consistent class with other words for familial relations, i.e. *bróðir* 'brother', *faðir* 'father', and *systir* 'sister', on both morphophonological and semantic grounds. In this connection, as discussed in Paper I (PI:15) and Paper II (PII:216), it is worthy of note that a strong semantic association with biological sex can impede the rate at which grammatical gender is projected onto an inflectional form that otherwise exhibits definitive phonetic coherence with the schema(s) for a given prototype (see below).

As noted in 3.2.2, schemas are cognitive points of reference that abstract over the formal and functional attributes of a set of items perceived as similar. The process of schematisation takes its cue from factors such as common phonetic structure, semantics, grammatical gender, appropriate contexts of use etc. (following Bybee 2001:27). Therefore, in light of the fact that 91.89% of nouns in plural *-ur* are feminine, the schemas $[Xur]_{\text{nom./acc.pl.}}$ and $[XæTur]_{\text{nom./acc.pl.}}$ are necessarily imbued with cue validity for feminine grammatical gender. However, association of inflectional forms in plural *-ur* with human referents of male *biological sex* may hinder reanalysis even when the form in question aligns with the relevant schema (PI:15; PII:216).

Significantly in this connection, despite the fact that the phonetic structure of masculine plural ***brœður*** 'brothers' is highly similar to that of plural ***mæður***, the token frequency of innovative fem.pl.def. *brœðurnar* indicates that plural *brœður* is very rarely attracted by the schema $[XæTur]_{\text{nom./acc.pl.}}$. Indeed, plural *brœður* and *feður* 'fathers', unlike e.g. plural *fætur*, are only rarely — if at all — reanalysed as feminine (PI:14). In Paper I, this dynamic is attributed to a strong semantic association of *bróðir* and *faðir* with male biological sex, i.e.

[54] Knudsen (1967) posits Ice. *bríx*, *flík*, *tík*, *vík* as a feminine microclass on phonetic grounds, e.g. due to the rhyme sequence *-ík*. Given that the microclass forms a phonetically well-defined subset in plural *-ur*, its schemas likely constitute another subordinate-level category within the broader feminine subtype. See Jónsdóttir (2020:27), who discusses the inflection of multisyllabic words (borrowings of Greek and Latin origin) in *-ík*, some of which get gen.sg. *-ur* and/or *-ar*.

they refer to men only. Therefore, in those instances where reanalysis of plural *bræður* and *feður* does occur, it is arguably semantic and/or formal association to feminine plural *dætur*, *mæður* and *systur* 'sisters' — all of which arguably constitutes a semantic class — that facilitates reanalysis as feminine, rather than alignment with the schema [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$[55]

By the same token, Paper II (PII:216) argues that assignment of feminine grammatical gender to the entire paradigm of Ice. *blók* is attributable to the fact that its use is not constrained by the biological sex and/nor gender identity of the referent. However, pragmatics plays a role in the assignment of feminine grammatical gender to forms of Ice. *kók*. Thus, the noun is almost always qualified by the feminine forms of modifiers in the count noun sense, i.e. in reference to a portion of the liquid delimited in contexts such as 'a bottle/can of coke'. Importantly, Icelandic *flaska* 'bottle' and *dós* 'can' are both grammatically feminine. Thus, it can be argued that in the cases of both *blók* and *kók*, formal alignment with schemas for the X*ó*/*æ*T-microclass is further assisted by semantics and pragmatics, i.e. the very factors argued to hinder reanalysis of plural *bræður* and *feður* as grammatically feminine.

In light of the above, phonetic coherence with a schema can be considered the main facilitating factor in both alignment and reanalysis of an inflectional form — all things being equal. However, it is clear that semantic and pragmatic factors also play a facilitating role. A combination of these factors likely accounts for the fact that plural *fætur* is reanalysed as grammatically feminine at a higher rate than any other masculine form in plural *-ur*, albeit in the minority of cases: plural *fætur* aligns perfectly with the prototype for the X*ó*/*æ*T-microclass in terms of phonetic coherence, cf. [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, while its referents are neutral with regard to biological sex and/or gender identity (see PI:14, Table 1).

5.3 Paper I: Reanalysis of masculine forms in plural -ur as feminine

The current section fleshes out the analysis presented in Paper I, accounting for the linguistic and non-linguistic factors that facilitate the reanalysis of masculine forms in plural *-ur* as feminine at different rates of frequency. Subsection 5.3.1 demonstrates that frequency of use impacts categorisation via the domain-general cognitive process of statistical learning, discussed in 3.2.2. In subsection 5.3.2, I elaborate on cognitive economy as a property of functionally related linguistic categories, as discussed in the previous subsection. Subsection

[55] Reanalysis of plural *bændur* 'farmers' as feminine at a higher rate than either plural *bræður* or *feður* is perhaps due to the fact that the word *bóndi* 'farmer' can — perhaps more now than before — apply to people of both sexes (see PI:15).

5.3.3 presents evidence from actual language use in support of hierarchical relations between varyingly schematic representations of linguistic categories. In subsection 5.3.4, I discuss the effects of minimal schematicity as a facilitator of limited productivity in light of arguments forwarded in the preceding subsections. Subsection 5.3.5 models the extent to which varying degrees of schematicity can facilitate different rates of productivity for respective schemas represented at distinct levels of abstraction. This model is characterised as the 'net effect'. In subsection 5.3.6, I present a summary of the current section.

5.3.1 Categorisation as a function of statistical learning

In Section 5.2, I delineated the prototype structure of the Icelandic feminine subtype in plural *-ur*, as it centres around the X*ó*/*æ*T-microclass. The posited structure of the subtype, it was argued, accords with Rosch's principle of cognitive economy in terms of the taxonomic arrangement of constituent schemas at distinct levels of abstraction. In the current section, I demonstrate that frequency of use impacts categorisation via the domain-general cognitive process of statistical learning, discussed in 3.2.2. Further, I demonstrate that via this process, we avail ourselves of tacit knowledge pertaining to the skewed frequencies of non-randomly co-occurring attributes which characterise linguistic and non-linguistic structures discerned in our environment (e.g. Taylor 2012:187). Moreover, due to the impact of co-occurring attributes on categorisation, statistical learning is considered to directly facilitate the function of categorisation, i.e. to reduce uncertainty in our environment.

To demonstrate the influence of statistical learning on categorisation, let us consider an exceptionless and, therefore, prototypical attribute of feminine and neuter nouns in Icelandic, the schema for which abstracts over the rule of referral in (20).

(20) The accusative plural form of all feminine and neuter nouns is the same as the nominative plural.

It is important to note that the stipulation in (20) applies to literally thousands of paradigms in Icelandic, meaning that the relevant schema is entrenched at a high level of abstraction across all relevant instances of prior experience.[56] However, given that Icelandic has three grammatical genders, mention of two of these in (20) implies that the stipulation is not generally applicable to masculines. Indeed, this is correct, cf. that alternation of the kind

[56] In fact, the stipulation in (20) applies to the paradigms of *all* nominals, i.e. nouns, pronouns, adjectives, numerals, in the nominative/accusative plural feminine and neuter forms.

nom.pl. *hest**ar*** ~ acc.pl. *hest**a*** ‘horses’ and *gest**ir*** ~ *gest**i*** ‘guests’ is applied in the inflection of the majority of Icelandic masculine nouns (see the relevant discussion around the stipulations stated in (11) and (12) in 3.2.2).

Despite this fact, adherence to the pattern of alternation nom.pl. *-ar* ~ acc.pl. *-a* or *-ir* ~ *-i* is not a prerequisite for assignment of masculine grammatical gender, as demonstrated by the relatively few masculine nouns in plural *-ur*, e.g. *bænd**ur*** ‘farmers’, *eigend**ur*** ‘owners’, *fæt**ur***. Crucially, 91.89% of nouns in plural *-ur* are feminine. Therefore, there is a degree of crossover between the stipulation in (20) — which applies generally to feminines and neuters only — and that which applies to the inflection of masculines and feminines in plural *-ur*, cf. the rule of referral in (21).

(21) A nominative plural in *-ur* has the same form as the accusative plural.

Note that (21) makes no reference to grammatical gender. However, as argued by Bybee (2001:27), schemas necessarily contain *all* information pertinent to the items over which they abstract. In other words, we can assume that the knowledge inherent to (21) is informed by statistical learning, the process by which knowledge of the dispersion of plural *-ur* and, therefore, its cue validity for assignment of feminine grammatical gender is amassed. Given the above, what answers can we provide to the question posed in (22), pertaining to the function that Rosch attributes to categorisation?

(22) How does reference to the stipulations in (20) and (21) facilitate uncertainty reduction?

To answer this question, let us first elaborate on this function in domain-general terms:

> [T]o treat [the] segmentation of the world as originally arbitrary … would be reasonable only if the world were entirely unstructured; that is … if the world formed a set of stimuli in which all possible stimulus attributes occurred with equal probability combined with all other possible attributes. (Rosch 1976:383)

In other words, the world is not an unstructured entity. Rather, some phenomena are perceived as similar due to the non-randomly skewed frequencies with which distinctive sets of attributes co-occur. Indeed, it is similarity based on such skewed frequencies that facilitates assignment of items such as dining chairs, garden chairs, rocking chairs, armchairs,

and beanbags to the basic level category CHAIR, rather than, say, DOG, despite the fact that members of both categories typically have four "legs". Further, it is the varyingly consistent combination of co-occurring attributes such as four wooden legs, a horizontal seat on top of those legs, and, typically, a vertical or almost vertical back attached to the seat that facilitates elaboration of the schema for the relevant basic level category.

Note that while rocking chairs, armchairs, and beanbags may not to have legs, they share other formal attributes with dining chairs and garden chairs. Further, all instantiations of the category CHAIR perform a continuous function — they allow people to take the weight off their feet by sitting. The non-random frequencies with which these formal attributes co-occur to different degrees across functionally related items informs representation of the category at the basic level. Further, Rosch and others appear to view categorisation as a function of the domain-general cognitive process of analogy, via which entrenched knowledge of the world facilitates structure mapping and subsequent alignment between relationally similar and functionally parallel sets of attributes.

In other words, the skewed frequencies at which non-randomly co-occurring attributes are integrated into separate phenomena determines the likelihood that a newly encountered item will be assigned to a given category. In turn, the degree to which the newly assigned item shares a random number of attributes with all existing members determines the status within the category, i.e. whether it is considered prototypical (think dining chair) or not (think bean bag) (see 3.3 on family resemblance structure within categories). As a function of this process, categorisation reduces uncertainty, both in terms of category content and in relation to category-external phenomena.

In order to demonstrate that linguistic categorisation satisfies the domain-general function of analogical reasoning, it is necessary to show that linguistic categories are also established and expanded through analogy. To this endeavour, Taylor (2012:193–194) argues that language exhibits structural qualities attributed by Rosch (1978:29) to the world more generally:

> Language, too, does not present itself to us as 'an unstructured set of equiprobable elements', whether the elements be sound segments, syllable types, words, or constructions. The elements that we perceive in a language do not occur uniformly and they are not distributed randomly over the speech that we encounter. It is the 'correlational structure' of language which makes possible the emergence of linguistic categories, thus ensuring both the viability of the language as a system of communication as well as its learnability. (Taylor 2012:193–194)

Taylor's argument is, therefore, valid to the extent that non-randomly skewed frequencies of co-occurring attributes have cue validity for assignment to linguistic categories also (see above). In this connection, it can be demonstrated that inflectional categories are characterised by non-randomly co-occurring sets of attributes, by which formal and functional relations between inflectional forms have cue validity for inflection class membership. In other words, as the vast majority of inflection classes in Icelandic contain nouns of a given gender, the nature of intra-paradigmatic relations between forms serves to reduce uncertainty with regard to gender assignment. Therefore, it also assists with agreement in the broader morphological and/or syntactic context, such as indicating the appropriate form of the definite article, or the morphologically appropriate marking of determiners.

In domain-general terms, given that non-randomly co-occurring attributes determine perceived structure in our environment, the tracking of skewed frequencies can be viewed as a necessary function of statistical learning. Proceeding on the assumption that language change is subject to constraints imposed by domain-general cognition on language use, it is claimed that knowledge pertaining to the skewed frequencies of co-occurring attributes that distinguish inflectional patterns is entrenched via statistical learning (see Jost and Christiansen 2017). In light of this claim, it proves necessary to demonstrate that the properties of inflection classes noted above also have cue validity for appropriate morphological treatment, i.e. gender assignment (form), and syntactic treatment, i.e. agreement (function).

Taylor (2012:187), based on Murphy (2002:215), argues that the *cue validity* of an inflectional form for class membership triggers the **category validity** of that form in terms of appropriate usage.[57] In other words, assignment of a noun to a given inflection class based on similarity to other class members (cue validity) facilitates further inferences as to how forms of the noun should be treated both within and beyond the context of the paradigm in terms of gender agreement (category validity). The principles that govern categorisation in language are, therefore, the same as those that facilitate assignment of four-legged items with a seat and upright back piece to the category CHAIR, the schema for which also abstracts over functional attributes such as facilitating rest.

[57] The term 'category validity' does not occur in Papers I–III. This is because both referees for Paper I advised that I collapse of distinction between *cue validity* and *category validity* for the sake of expediency. Therefore, I resolved to use the term 'cue validity' to convey both senses.

Let us now impose the theme of this discussion onto inflectional categories in Icelandic, specifically those alluded to — both explicitly and implicitly — by the stipulations in (20) and (21). For some nouns, instantiation of the plural nominative and accusative is characterised by formal distinction of the kind nom.pl. *-ar* ~ acc.pl. *-a* or *-ir* ~ *-i* (see above; also PI:11). As formal distinction of this kind is characteristic of the inflection of the vast majority of Icelandic masculine nouns and, crucially, occurs among masculines only, it is considered a prototypical attribute of their inflection. This view is supported by the fact that forms which engage in this relation are hardly ever reanalysed as feminine. Thus, as a function of prototypicality, the non-randomly co-occurring cluster of attributes that defines relations of the kind *-ar* ~ acc.pl. *-a* or *-ir* ~ *-i* has high cue validity for the assignment of masculine grammatical gender.

Conversely, Icelandic masculine forms in plural *-ur*, e.g. masculine *fæt**ur***, are occasionally treated as feminine, as betrayed by their occurrence with the feminine form of the article, cf. pl.**def**. *fætur**nar***, instead of masculine nom.pl.**def**. ***-nir***, acc.pl.**-def**. ***-na*** (see PI:13–14 on different rates of reanalysis). Expressed in terms of cue and category validity, the skewed dispersion of plural *-ur* among masculine and feminine nouns, particularly in relative juxtaposition to the high frequency of the relations nom.pl. *-ar* or *-ir* and acc.pl. *-a* or *-i* among masculines only, means plural *-ur* has high cue validity for feminine grammatical gender.[58] In turn, categorisation as feminine motivates use as such.

In terms of prototype structure, despite the fact that forms such as plural *fætur* are used as masculine in the majority of instances, it can be claimed that they occupy a more peripheral space within the relevant category than do those in nom.pl. *-ar* or *-ir* and acc.pl. *-a* or *-i*. In other words, masculine forms in plural *-ur* are best characterised as satellites, occupying border space with the equivalent category for feminine forms, in which the ending plural *-ur* has more central status.

Let us, then, attempt an answer to the question posed in (22), above. To this end, it is important to distinguish between language use as *prescribed* and language use as *described*. In other words, if we equate certainty with the "correct" use of language as prescribed by school grammars and the like, characterised as e.g. "The plural nominative/accusative for

[58] Markússon (2023b) demonstrates that the same factors motivate use of Far. masculine plural *føtur* 'feet' as feminine. Further, he shows that the skewed dispersion of the Faroese plural ending *-ar* among feminine and masculine nouns — overwhelmingly in favour of the masculines — is responsible for the reanalysis of feminine nouns in this ending as masculine. This is taken to suggest that dispersion perhaps determines the rate of gender reanalysis in the broader context.

fætur is masculine and, therefore, should not occur with the feminine form of the article", the answer to (22) is: The stipulations in (20) and (21) do not reduce uncertainty but rather facilitate the incorrect use of language.

However, we get a different answer if we concede that uncertainty arises due to the exceptionless occurrence of syncretism among the literally thousands of Icelandic feminines, as well as properties of the dispersion and resultant peripheral position of plural *-ur* within the relevant category for masculines. In other words, those instances in which historically masculine forms in plural *-ur* occur with the feminine definite article, e.g. *bændur**nar***, *eigendur**nar***, *fætur**nar***, likely reflect momentary uncertainty with regard to appropriate use on the very basis of the stipulations in (20) and (21). Therefore, outcomes in the feminine article *-nar* are a reasonable attempt at uncertainty reduction, i.e. the main function of categorisation, betraying reliance on the skewed dispersion of plural *-ur* among masculine and feminine classes. In light of this answer, the analysis presented in Paper I demonstrates that the prototype structure of categories is a function of statistical learning and, thus, supports a view of rich memory for language.

5.3.2 Elaborating on taxonomies of increasing schematicity

In the current subsection, I elaborate on cognitive economy as a property of functionally related linguistic categories, as discussed in 5.3.1. Following e.g. Audring 2019, Paper I argues for and posits a taxonomy of formally distinct but functionally continuous schemas, characterised by a hierarchy of abstraction (see PI:12). Each level of the taxonomy is represented by its own schema, which abstracts over inflectional forms in plural *-ur*. The hierarchical nature of the taxonomy is defined by the number of one-to-one correspondences between a masculine form in plural *-ur*, such as plural *fætur* 'feet', *vetur* 'winters', *fingur* 'finger', *eigendur* 'owners', *bræður* 'brothers', and the functional attributes conveyed by a given schema. It is argued that the extent of alignment between form and schema determines the rate that a given masculine form in plural *-ur* is reanalysed as feminine — all other things being equal.

The standard paradigms for Ice. masculine *fótur* 'foot', *vetur* 'winter', *fingur* 'finger', and *eigandi* 'owner' are shown in (23), with the article for the nominative and accusative plural forms in **bold**.

(23)					
	sg. nom.	*fótur*	*vetur*	*fingur*	*eigandi*
	acc.	*fót*	*vetur*	*fingur*	*eiganda*
	dat.	*fæti*	*vetri*	*fingri*	*eiganda*
	gen.	*fótar*	*vetrar*	*fingurs*	*eiganda*
	pl. nom.(-def.)	*fætur**nir***	*veturnir*	*fingurnir*	*eigendurnir*
	acc.(-def.)	*fætur**na***	*veturna*	*fingurna*	*eigendurna*
	dat.	*fótum*	*vetrum*	*fingrum*	*eigendum*
	gen.	*fóta*	*vetra*	*fingra*	*eigenda*

As highlighted in 5.1, masculine plural *vetur*, *fingur*, and *fætur* serve as the respective intra-paradigmatic sources for fem.pl.def. *veturnar*, *fingurnar*, and *fæturnar*. The overtly feminine definite forms, it is argued, occur when the former set is reanalysed as feminine, i.e. subsequent to alignment with the basic level schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$, which has a 91.89% dispersion rate among feminine nouns.

In Paper I, it is argued that language users occasionally analyse the morphological composition of inflectional forms such as plural *fætur*, on the one hand, and plural *vetur*, *fingur*, *bræður*, on the other, in a way contrary to etymological regularity. The general view is that the sequence -(*u*)*r* in *vetur* and *fingur* actually belongs to the stem (see below).[59] In this connection, the vast majority of strong masculine nouns in nom.sg. *-ur* have no ending in the accusative singular, cf. nom.sg. *fót**ur*** ~ acc.sg. *fót*, also e.g. masc. *hest**ur*** ~ *hest*, *gest**ur*** ~ *gest*. This morphological attribute is taken to indicate that *-ur* is an inflectional ending in such paradigms (Kvaran 2005:242, 304; see PI:10–11).

Conversely, the sequence *-ur* is present in the syncretic nominative/accusative singular of e.g. *vetur* and *fingur*, as well as in all other forms of these words. This fact lends itself — quite rightly in my view — to interpretation of *-ur* in e.g. nom./acc.sg., plural *vet**ur***, *fing**ur***, and *-r-* in e.g. dat.sg. *vet**r**i*, *fing**r**i* as part of the stem both diachronically and synchronically. However, below it is argued that the sequence *-ur* in e.g. plural *vetur*, *fingur*, *bræður* is occasionally analysed as an inflectional ending that is both formally and functionally identical to the same sequence in plural *fæt**ur*** (and e.g. *eigend**ur***). Given the structural and

[59] The sequence *-ur* in e.g. *vet**ur***, *fing**ur***, *bræð**ur*** alternates with *-r-* in forms whose ending is or begins with a vowel, cf. nom./acc.sg. *vet**ur*** but gen.pl. *vet**r**a*. For convenience, this element will henceforth be referred to as *-ur*.

etymological position endorsed above, what, then, might be the motivation for occasionally attributing the status of ending to the sequence in plural *vet**ur**, fing**ur**, brœð**ur***?

It is argued that an all-or-nothing structural approach proves unhelpful in accounting for the fact that masculine forms in plural *-ur* — be that sequence an inflectional ending or part of the stem etymologically — are occasionally reanalysed as feminine, while other forms of the paradigm are not. In light of this, any claim that analysis of plural *-ur* as an inflectional ending lacks intra-paradigmatic support can be countered through reference to the reanalysis of e.g. both *vetur* and *fætur* as feminine, which occurs *despite* intra-paradigmatic alternation with forms that are only ever used as masculine. Thus, we get the relations masc.nom./acc.sg.def. *veturinn* ~ ?fem.pl. *vetur* ~ fem.pl.def. *vetur**nar***; masc.nom.sg. *fótur* ~ masc.dat.sg. *fæti* ~ ?fem.pl. *fætur* ~ fem.pl.def. *fætur**nar***. As noted in Paper I, the rarity or non-occurrence of feminine singular forms of *fótur* likely stems from the status of nom.sg. *-ur* as an indicator of masculine grammatical gender due to properties of its dispersion. Why, then, do e.g. plural *vetur* and *fætur* serve as intra-paradigmatic sources of fem.pl.def. *veturnar*, *fæturnar* when the former pair occupies their respective paradigms with masculine-only forms?

Paper I argues that this is due to formal identity between the historically stem-final sequence *-ur* in plural *vet**ur*** (also e.g. plural *fing**ur*** and *brœð**ur***) and the inflectional ending plural *-ur* (PI:10–11). As already noted, plural *-ur* has strong cue validity for assignment of feminine grammatical gender on account of its dispersion. Further, as also argued in Paper I, the semantically significant difference between singular and plural in nouns is reflected in the reanalysis of e.g. plural *vetur* as feminine, despite the fact that the same process never affects singular forms of the word. In other words, reanalysis occurs as a two-step process that affects individual forms rather than all members of the paradigm simultaneously (see Chapter 4). Thus, for the reasons just stated, it is concluded that a lack of support from other members of the paradigm does not suffice to prevent reanalysis of e.g. plural *vetur* as feminine on account of the ”ending” plural *-ur*.

While 8.11% of masculine nouns shows the historically regular inflectional ending plural *-ur*, most of these belong to the class of *nd*-stems, all of which are masculine (see Iversen 1972:66; also PI:12–15). The class in question conforms to the intra-paradigmatic relation $[\mathrm{X\underline{A}}nd\boldsymbol{i}]_{\text{nom.sg.}}$ ~ $[\mathrm{X\underline{A}}nd\boldsymbol{a}]_{\text{acc./dat./gen.pl.}}$ ~ $[\mathrm{X\underline{E}}nd\boldsymbol{ur}]_{\text{nom./acc.pl.}}$, e.g. nom.sg. *eig**a**ndi* ~

acc./dat./gen.sg *eigandа* ~ pl. *eigendur*.[60] It is argued that the very specific association of the pattern sg. *-and-* ~ pl. *-end-* with the inflection of masculine nouns only is likely responsible for the relatively low frequency with which words containing it are reanalysed as feminine, all below 5.5% of instances (see PI:14, Table 1). Indeed, the schema $[Xur]_{\text{nom./acc.pl.}}$ should otherwise be schematic enough to attract masculine forms such as plural *eigendur*, which shares the suffix pl. *-end-*, as well as plural *-ur*, with a microclass of feminines, cf. e.g. sg. *strönd* ~ pl. *strendur* 'coastline, beach(es)' (following Booij 2010:41; Brown and Hippisley 2012:34).

As already intimated, actual usage occasionally betrays reanalysis of the nominative/accusative plural forms of *nd*-stems as feminine, cf. fem.pl.def. *eigendurnar* 'the owners' at a rate of 2.83%, *bændurnar* 'the farmers' at 2.87%, *nemendurnar* 'the students' at 1.62%, instead of original masc.nom.pl.def. *eigendurnir*, *bændurnir*, *nemendurnir* and acc.pl. *eigendurna*, *bændurna*, *nemendurna*. However, as reported in Paper I, the rate at which *nd*-stem sources in plural *-ur* are reanalysed as feminine is much lower than for plural *fætur*, *fingur*, and *vetur* (PI:14). Further, this lower rate of reanalysis can likely also be attributed in part to the fact that *nd*-stems have human referents and, therefore, often refer to men.

In light of the schematic and semantic factors considered to impact the rate of reanalysis for *nd*-forms in plural *-ur*, Paper I focuses particularly on the facts pertaining to the different rates of reanalysis reported for plural *fætur* at 22.82%, *fingur* at 15.12%, and *vetur* at 9.74% of instances, respectively (see PI:14, Table 1). The descending token frequencies of fem.pl.def. *fæturnar*, *veturnar* and *fingurnar* betray a mismatch in the rate of reanalysis relative to the individual token frequencies of original/standard masc.nom.pl.def. *fæturnir*, *veturnir*, *fingurnir* and masc.acc.pl.def. *fæturna*, *fingurna*, *veturna*. Appeal is made to the varying degrees of schematicity that characterise the taxonomy referred to above. In order to provide a convincing account of the effects of schematicity on the different rates of reanalysis reported, a more detailed word on the dispersion of plural *-ur* among masculine and feminine classes is in order.

As discussed in 3.2.2 and 5.3.1, the endings for the plural nominative and accusative forms of the vast majority of Icelandic masculine nouns can be represented schematically as $[XV_1r]_{\text{nom.pl.}}$ and $[XV_1]_{\text{acc.pl.}}$, respectively. In the context of intra-paradigmatic alternation, these schemas constitute the sister schema $[XV_1r]_{\text{nom.pl.}} \sim [XV_1]_{\text{acc.pl.}}$, cf. nom.pl. *-ar* ~ acc.pl.

[60] Here, 'X' refers to the root sequence of such words, e.g. *eig-* in ***eig**-and-i*, *nem-* in ***nem**-and-i* 'student'. The notation $[...\underline{A}...]_{\text{sg.}} \sim [...\underline{E}...]$ generalises over the fact that, while the vast majority of *nd*-stems inflects like *eigandi*, cf. suffixes in sg. *-a-*, pl. *-e-*, a small number does not. Thus, Ice. *bóndi* engages in alternation between sg. *-ó-* and pl. *-æ-*, cf. nom.sg. *bóndi* ~ acc./dat./gen.sg. *bónda* ~ nom./acc.pl. *bændur*.

-a, e.g. *hest**ar*** ~ *hest**a***; nom.pl. *-ir* ~ acc.pl. *-i*, e.g. *gest**ir*** ~ *gest**i***. The dispersion of the sister schema is associated with masculine classes only and is of medium type frequency in Icelandic (based on statistical data collated by Svavarsdóttir 1993). However, the functionally equivalent forms of some masculine nouns, e.g. plural *fætur*, *eigendur*, demonstrate that conformity to this relation is not a prerequisite for categorisation and use as masculine. Indeed, masculine forms in plural *-ur* most often alternate with overtly masculine plural definite forms, cf. masc.nom.pl.**def**. *fætur**nir***, *eigendur**nir*** and masc.acc.pl.def. *fætur**na***, *eigendur**na***, i.e. these are more common than their functionally equivalent feminine doublets.

Conversely, syncretism in nominative/accusative plural is an exceptionless formal attribute of feminine (and neuter) nouns, e.g. feminine plural *myndir* ‘pictures, forms’, *greinar* ‘branches, articles’, *stelpur* ‘girls’, *mýs* ‘mice’. In (24), the paradigms of feminine *mynd*, *grein*, *stelpa* and *rót* are given alongside the plural forms of masculine *fótur* (the endings of the nominative/accusative plural are in **bold**).

(24)	sg.	nom.	*mynd*	*grein*	*stelpa*	*rót*	
		acc.	*mynd*	*grein*	*stelpu*	*rót*	
		dat.	*mynd*	*grein*	*stelpu*	*rót*	
		gen.	*myndar*	*greinar*	*stelpu*	*rótar*	
	pl. nom./acc.		*mynd**ir***	*grein**ar***	*stelp**ur***	*ræt**ur***	*fæt**ur***
		dat.	*myndum*	*greinum*	*stelpum*	*rótum*	*fótum*
		gen.	*mynda*	*greina*	*stelpna*	*róta*	*fóta*

As noted in Paper I (PI:11), entrenched knowledge of this pervasive and characteristic formal distinction between masculine and feminine classes is necessarily stored at a high level of abstraction (see e.g. Janda 2002; 2007; also below). However, at the physical level of language use, the distinction is instantiated as in the masculine forms in (25a), on the one hand, and the feminine forms in (25b), on the other. The relevant plural forms of masculine *fótur*, *vetur* and *fingur*, which straddle the masculine-feminine border, are given in (25c).

(25)	a.	$[X\mathbf{V_1r}]_{\text{nom.pl.}}$ ~	$[X\mathbf{V_1}]_{\text{acc.pl.}}$	(masc.) *hest**ar*** ~ *hest**a***, *gest**ir*** ~ *gest**i***
	b.	$[X\mathbf{Vr}]_{\text{nom./acc.pl.}}$		(fem.) *mynd**ir***, *grein**ar***, *stelp**ur***, *ræt**ur***
	c.	$[X\mathbf{Vr}]_{\text{nom./acc.pl.}}$		(masc./?fem.) *fæt**ur***, *vet**ur***, *fing**ur***

(Lifted from PI:12)

Given the view that schemas are not elaborated independently of the formal, functional, and interactional attributes over which they abstract (e.g. Bybee 2001:27; Lakoff 2018:86–87), the notation '-V_1-' should be taken to imply the phonologically arbitrary subset {a, i} only. In other words, statistical learning yields knowledge that only the relations nom.pl. *-ar* ~ acc.pl. *-a* and *-ir* ~ *-i* can instantiate the sister schema $[XV_1r]_{\text{nom.pl.}}$ ~ $[XV_1]_{\text{acc.pl.}}$.

Conversely, highly schematic $[XVr]_{\text{nom./acc.pl.}}$ in (25b), whose dispersion is overwhelmingly associated with feminine classes, offers a tried and tested point of reference for successful deduction of forms in nom.pl. *-ir*, *-ur* and *-ar* from acc.pl. *-ir*, *-ur*, and *-ar* and vice versa. Therefore, this property of the schema's dispersion greatly reduces the likelihood that it should project masculine gender onto the inflectional forms referred to it. On the contrary, indeed, Paper I argues that we should *expect* alignment of e.g. plural *fætur* with the schema $[XVr]_{\text{nom./acc.pl.}}$ to facilitate reanalysis as feminine. With this in mind, let us now consider the schemas to which forms in plural *-ur* might be referred for the purpose of categorisation.

The schema $[XæTur]_{\text{nom./acc.pl.}}$ is minimally schematic, i.e. it renders highly specific formal and functional constraints imposed on the set of forms over which it abstracts. For this reason, it is considered subordinate to medially schematic $[Xur]_{\text{nom./acc.pl.}}$. In Audring's (2019) terms, the former is a daughter of the latter. In turn, the schema $[Xur]_{\text{nom./acc.pl.}}$ is posited as basic, i.e. the daughter of highly schematic $[XVr]_{\text{nom./acc.pl.}}$. This view is congruent with arguments expressed by e.g. Audring (2019), Barðdal (2008), Booij (2010) and Bybee (2001:8), and similar to Albright's (e.g. 2002; 2008; 2009) conception of form-to-form mapping "rules" of varying specificity (also Albright and Hayes 2003).

Thus, the hierarchical subordination of basic $[Xur]_{\text{nom./acc.pl.}}$ and subordinate $[XæTur]_{\text{nom./acc.pl.}}$ to superordinate $[XVr]_{\text{nom./acc.pl.}}$ is rendered as the taxonomy in (26).

(26) a. Superordinate $[X\mathbf{Vr}]_{\text{nom./acc.pl.}}$ *mynd**ir**, grein**ar**, stelp**ur**, rætur, fætur*
b. Basic $[X\mathbf{ur}]_{\text{nom./acc.pl.}}$ *stelp**ur**, vet**ur**, rætur, fætur*
c. Subordinate $[X\underline{æ}\underline{T}\mathbf{ur}]_{\text{nom./acc.pl.}}$ *r**æt**ur**, f**æt**ur***
(Based on PI:12)

Regarding the representation of inflectional endings in (26), the notation $[-Vr]_{\text{nom./acc.pl.}}$ in (26a) abstracts over syncretism in plural *-ar*, *-ir* and *-ur*, thus encompassing the ending represented in both $[Xur]_{\text{nom./acc.pl.}}$ and $[XæTur]_{\text{nom./acc.pl.}}$.

Thus, [X*ur*]$_{\text{nom./acc.pl.}}$ is neither schematic enough to abstract over the corresponding forms of all feminine nouns, cf. e.g. plural *mynd**ir***, *grein**ar***, nor is it too specific as to only facilitate assignment of a noun to the X*ó*/*æ*T-microclass. In this sense, it occupies "[t]he level used for everyday neutral reference" (Croft and Cruse 2004:83; see discussion in 3.2.4). Its status as a basic level category accounts for its function as a single point of reference for categorisation of non-feminine forms in plural *-ur* that are otherwise formally dissimilar (see the examples in (26b)). By the same token, [XV*r*]$_{\text{nom./acc.pl.}}$ is considered superordinate to [X*ur*]$_{\text{nom./acc.pl.}}$, as the former has "…fewer defining attributes than…" the latter (Croft and Cruse 2004:84). Therefore, [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ reflects a subordinate level category on account of the fact that "…members have high mutual resemblance…" (Croft and Cruse 2004:85). The hierarchical nature of the posited linguistic taxonomy is, therefore, congruent with Rosch's (1975) principle of cognitive economy.

5.3.3 Language use and hierarchies in linguistic categories

Positing the taxonomy in (26) is further justified by the fact that use of historically masculine plural *fætur* has shown development that is both convergent with and divergent from that of e.g. masculine plural *vetur* and *fingur*. In Paper I, it is argued that all three are reanalysed as feminine on account of syncretism in plural *-ur* — all other things being equal. In the current subsection, I focus on the different rates of reanalysis for plural *fætur*, *fingur*, and *vetur*, as the rates for *nd*-stems are largely insignificant — at 5.47% or less — likely due to association of *nd*-stem inflection with male referents (see discussion in 5.3.2). In light of this view, it is argued that the respective rates of reanalysis for plural *fætur*, *fingur*, and *vetur* speak specifically to the influence of varying degrees of schematicity, when all other things are equal.

As noted in Chapter 4, acc.pl.def. *veturna* is the most frequent of the relevant masculine forms, returning 6,913 results from the *IsTenTen* corpus. Next most frequent is masc.acc.pl.def. *fæturna*, with 2,882 results. Finally, a search for masc.nom.pl.def. *veturnir* returned 245 results, making it the least-frequent of the relevant masculine forms. Of their feminine doublets, a search of the same corpus returned more results for pl.def. *fæturnar* than for pl.def. *veturnar* and *fingurnar* combined, with 1,274, 773, and 285 results, respectively (see PI:14, Table 1).

As argued in Paper I (PI:14–16), the higher rate of reanalysis for plural *fingur* as feminine relative to *vetur* can be accounted for by a combined appeal to two factors. First, to the cue validity that the schema [X*ur*]$_{\text{nom./acc.pl.}}$ has for feminine grammatical gender.

Secondly, to semantic association of *fingur* with feminine forms such as plural *hendur* 'hands', (plurale tantum) *herðar* 'shoulders', even *tær* 'toes', and *lappir* 'legs', cf. pl.def. *hendurnar*, *herðarnar*, *tærnar*, and *lappirnar*. In this connection, it must be noted that semantic association with these same feminine forms likely also accounts in part for the reanalysis of plural *fætur* as feminine.

However, Paper I takes the much higher rate of reanalysis for plural *fætur* compared with *fingur* to suggest unequal semantic links with the feminine nouns for body parts listed above, meaning that all things might not be equal. Note that a common synonym of Ice. *fótur*, i.e. Ice. *löpp*, is feminine, while that of *fingur*, cf. Ice. *putti*, is masculine. Returning to the more general picture, the statistical information presented above demonstrates that the mismatch in respective rates of reanalysis for masculine plural *fætur* and *vetur* as feminine is significantly greater than those for *fingur*, on the one hand, and either *fætur* or *vetur*, on the other. For this reason, reflecting the path taken in Paper I, the analysis presented in the next subsection will focus mainly on the mismatched descending token frequencies of masculine and feminine doublets containing plural *fætur* and *vetur*. Despite this focus, reference will be made to other masculine forms in plural *-ur* where relevant.

5.3.4 Productivity: Types and degrees of similarity

This subsection elaborates on minimal schematicity as a facilitator of limited productivity. Further, I discuss the impact of degrees of similarity on the rate that masculine forms in plural *-ur* undergo reanalysis as feminine. In this connection, consider the question posed in (27), repeated from Paper I (PI:17).

(27) How do we account for the frequency relation between fem.nom./acc.pl.def. *fæturnar* and *veturnar*, on the one hand, and masc.nom.pl.def. *fæturnir*, acc.pl.def. *fæturna* and *veturnir*, *veturna*, on the other?

Answers to this question elucidate the impact of schematicity on the rate of reanalysis — all other things being equal.

With the question in (27) in mind, I endeavour below to characterise limited productivity in terms of categorisation, i.e. the domain-general cognitive process that attributes structure to phenomena in our environment. In this connection, deduction of the targets fem.pl.def. *fæturnar* and *veturnar* are considered varyingly specific consequences of the gang effect, different functions of which are properties of the prototype structure

delineated in 5.2 for the feminine subtype in plural *-ur*. The view taken in Paper I is that an appeal to alignment, reanalysis, and extension, as delineated by the proportional equations employed for exposition in historical linguistics, should not be allowed to speak for itself by rendering the *outcome* of analogical change only. Rather, Paper I seeks to render these underlying cognitive operations as the motivators and/or facilitators of change through innovative notation of proportions.

To this end, an understanding of two different kinds of similarity — each of which motivates analogy — is key. The type of similarity that exists between e.g. plural *rætur, kindur*, *fætur*, and *vetur* can best be characterised as **object similarity** (see e.g. Kotovsky and Gentner 1996:2798; see also Gentner 2005, who uses the term "overall similarity"). It is due to object similarity between separate instances of the common formal attribute *-ur*, itself an expression of the function *nominative/accusative plural*, that language users might perceive the four forms above as similar. Thus, object similarity based on plural *-ur* alone facilitates alignment with the basic level schema [X*ur*]$_{\text{nom./acc.pl.}}$. However, further points of similarity can be discerned between e.g. plural *rætur* and *fætur*, motivating alignment with subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, as well as the basic-level schema.

With regard to the degree of productivity exhibited by an inflection class, it has been demonstrated repeatedly that this property is particularly correlated with type frequency and schematicity. Thus, as noted several times, the interaction of high type frequency with high schematicity can facilitate all but limitless productivity. In relative terms, the schema [X*ur*]$_{\text{nom./acc.pl.}}$ is associated with medium type frequency in Icelandic (based on frequency counts by Svavarsdóttir 1993), while its schematicity is also within the medium range (see the discussion surrounding (26) in 5.3.2). Therefore, given this often-observed correlation, we might expect to attribute a higher rate of productivity to medially schematic [X*ur*]$_{\text{nom./acc.pl.}}$ than to minimally schematic [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ — all other things being equal.

As argued in Paper I (PI:11–12), any inflectional form in plural *-ur* can be aligned with medially schematic [X*ur*]$_{\text{nom./acc.pl.}}$, while only those that show the relevant additional points of similarity can align with the minimally schematic [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$. Thus, it is argued, different rates of reanalysis for masculine forms in plural *-ur* as feminine are a function of the degree of formal and functional one-to-one correspondence between a schema and an inflectional form that has become distanced — albeit only momentarily — from other members of its paradigm due to formal ambiguity. Once the form has been aligned with the relevant schema on formal and functional grounds, projection of morphological content from the latter to the former commences. Thus, Paper I argues, subsequent to alignment, projection

of feminine grammatical gender is the initial function of one-to-one correspondence, as depicted in (28) (see below on the notation employed).

(28) $[X\textit{ur}]_{\text{nom./acc.pl.}}$

nom./acc.pl. *stelpur*

↓

nom./acc.pl. *vetur*

The schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$ is included in (28) to represent the knowledge of form and function that is projected onto plural *vetur* (or any functionally equivalent masculine form in plural *-ur* that might be substituted for it). Projection of the morphological content, i.e. the attribute *feminine grammatical gender*, is represented by the arrow '↓'. Thus, (28) depicts the projection of functional content from feminine forms that align with the schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$ onto plural *vetur*.

Subsequent to projection, once a form has been categorised as feminine, it is highly likely to have category validity for further alignment with its sister schema $[X\textit{urnar}]_{\text{nom./acc.pl.def.}}$ due to the high rate of dispersion that the schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$ has across feminine paradigms. Therefore, it is this property that subsequently activates the former schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$ as an **output schema**, i.e. one which has category validity for a given pattern of intra-paradigmatic alternation, in this case according to the sister schema $[X\textit{ur}]_{\text{nom./acc.pl.}} \sim [X\textit{urnar}]_{\text{nom./acc.pl.def.}}$. Thus, reanalysis can be posited as the second function of one-to-one correspondence, while deduction of the target form is the third.

Perception of object similarity across distinct intra-paradigmatic relations is likewise a prerequisite for the deduction of a target, i.e. the D-form of a proportional equation. However, it is the perception of similarity *among* relations that informs this process. For this reason, the second kind of similarity to be discussed is best characterised as **relational similarity** (e.g. Kotovsky and Gentner 1996:2798). Thus, alternation between e.g. fem.pl. *stelpur*, *rætur*, and fem.pl.def. *stelpurnar*, *ræturnar*, respectively, assists in establishing parallel one-to-one correspondences across relations that already align with the sister schema $[X\textit{ur}]_{\text{nom./acc.pl.}} \sim [X\textit{urnar}]_{\text{nom./acc.pl.def.}}$. In turn, a form in plural *-ur* that is aligned with the output schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$ as a means of uncertainty reduction, i.e. the function of categorisation, subsequently serves as the source for alignment with overtly feminine $[X\textit{urnar}]_{\text{nom./acc.pl.-def.}}$. Therefore, the entrenched knowledge represented by sister schemas

facilitates parallel connectivity, by which two relations and their arguments are placed in correspondence according to role.

In other words, the relational knowledge that defines a sister schema is projected in order to extend an inflectional pattern by attracting a target via the category validity of the source. Further, it has been demonstrated that parallel connectivity is the prerequisite of **structural alignment**, i.e. successful alignment between relations, which is considered the hallmark of analogy (e.g. Gentner and Markman 1997; Kotovsky and Gentner 1996). An exemplary depiction of the process that yields structural alignment with the sister schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$ ~ $[X\textit{urnar}]_{\text{nom./acc.pl.def.}}$ is shown in (29).

(29)

$[X\textit{ur}]_{\text{nom./acc.pl.}}$	$\rightarrow$	$[X\textit{urnar}]_{\text{nom./acc.pl.def.}}$
nom./acc.pl. *stelpur*	~	nom./acc.pl.def. *stelpurnar*
↓		
nom./acc.pl. *vetur*	$\rightarrow$	nom./acc.pl.def. X; X = *veturnar*

To delineate in terms of the proportional schema A : B :: C : D, (29) implies that object similarity between the A-form, i.e. *stelpur*, and the C-form, *vetur*, as predicated on the common ending plural *-ur*, facilitates alignment with the schema $[X\textit{ur}]_{\text{nom./acc.pl.}}$. Therefore, alignment on formal grounds assists in establishing one-to-one correspondence that pertains to both form and function. Subsequently, one-to-one correspondence facilitates projection of morphological content onto the C-form, (now feminine) plural *vetur*. Intra-paradigmatic alternation as instantiated by the A- and B-forms in (29), i.e. plural *stelpur* and pl.def. *stelpurnar*, respectively, activates $[X\textit{ur}]_{\text{nom./acc.pl.}}$ with category validity for intra-paradigmatic alternation with its sister schema $[X\textit{urnar}]_{\text{nom./acc.pl.def.}}$. This property of the former schema is represented by the symbol '$\rightarrow$'. Once the morphological content of $[X\textit{ur}]_{\text{nom./acc.pl.}}$ has been projected onto the C-form, i.e. plural *vetur*, it is likewise imbued with category validity for intra-paradigmatic alternation of the kind $[X\textit{ur}]_{\text{nom./acc.pl.}}$ ~ $[X\textit{urnar}]_{\text{nom./acc.pl.def.}}$. Thus, the only logical solution to (29) is fem.pl.def. *veturnar*.

Due to this function of analogical reasoning — as assisted by the domain-general cognitive process of statistical learning — functionally equivalent forms of Ice. *fótur* could easily be substituted for the respective forms of *vetur* in (29). In that case, the D-form of the relevant proportion would logically be solved as fem.pl.def. *fæturnar*, based on the C-form, (now feminine) plural *fætur*. In line with arguments forwarded in Paper I, the following

subsections argue that masculine plural *fætur* aligns with the formal and functional attributes associated with both basic [X***ur***]$_{\text{Nom./acc.pl.}}$ and subordinate [X*æ*T***ur***]$_{\text{nom./acc.pl.}}$ as a means of uncertainty reduction, i.e. when momentarily distanced from other members of the paradigm during a usage event. Alignment at both levels of abstraction is depicted by the proportion in (30).

(30)

	[X***ur***]$_{\text{nom./acc.pl.}}$	→	[X***urnar***]$_{\text{nom./acc.pl.def.}}$
	[X*æ*T***ur***]$_{\text{nom./acc.pl.}}$	→	[X*æ*T***urnar***]$_{\text{nom./acc.pl.def.}}$
	nom./acc.pl. *stelp**ur***	~	nom./acc.pl.def. *stelp**urnar***
	nom./acc.pl. *r**æt**ur***	~	nom./acc.pl.def. *r**æt**urnar***
	↓		
	nom./acc.pl. *f**æt**ur***	→	nom./acc.pl.def. X; X = *f**æt**urnar***

In (30), one-to-one correspondences based on the ending plural *-ur* are in bold, demonstrating that this is a formal and functional attribute common to the relevant forms at both levels of abstraction. However, instances of underlined *æ*T convey one-to-one correspondence that is specific to the subordinate level only.

Paper I argues that while e.g. plural *vetur* aligns perfectly at the basic level, it does so more tentatively than plural *fætur* at the subordinate level. Therefore, at this stage, the following answer to the question posed in (27) is proposed: plural *vetur* is more likely to escape alignment with subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ on schematic grounds. Further, though plural *vetur* were occasionally attracted at the subordinate level, it would at best be peripheral to the X*ó*/*æ*T-microclass and not assigned to any specific feminine class. Conversely, plural *fætur* is not only reanalysed as feminine on alignment with [X***ur***]$_{\text{Nom./acc.pl.}}$ but is also specifically categorised as a X*ó*/*æ*T-microclass form on alignment with subordinate [X*æ*T***ur***]$_{\text{nom./acc.pl.}}$. Therefore, plural *fætur* is considered less likely to escape reanalysis as feminine than is plural *vetur*. The next subsection delineates the means employed in Paper I of modelling varying degrees of schematicity and the impact of this factor on the rate of reanalysis.

5.3.5 The 'net effect'

Paper I employs the metaphor of the 'net effect' when modelling the extent to which varying degrees of schematicity facilitate different rates of productivity for respective schemas, each of which represents a distinct level of abstraction (PI:17–18). The **net effect of minimal**

schematicity is a *network model* interpretation of the gang effect, by which a network of connecting lines demonstrates the skewed frequencies of formal and/or functional attributes common to sets of forms. This approach to modelling productivity is based on Bybee's Network Model (e.g. 1985, 2001, 2010). The innovative notations employed show that the greater the degree of one-to-one correspondence across forms, the greater the cue validity that the relevant cluster of attributes will have for assignment of grammatical gender and, potentially, for assignment of a form to a specific inflection class. In turn, the network approach demonstrates that once gender is assigned on this basis, category validity for use according to that assignment is activated.

As noted in 5.3.4, Paper I accounts for the mismatch in descending token frequency between fem.pl.def. *veturnar* and *fæturnar*, on the one hand, and their respective masculine doublet forms, nom.pl.def. *veturnir*, acc.pl.def. *veturna* and *fætur-nir*, *fætur-na*, on the other. There, it is argued that alignment of plural *fætur* with respective schemas that represent distinct levels of abstraction facilitates the relatively general net effect of basic [X*ur*]$_{\text{nom./acc.pl.}}$, while alignment with subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ nets plural *fætur* and potentially some instances of peripheral plural *vetur* closer to the physical level of language use. Thus, as alluded to in 5.3.4, the net effect can be expressed as the rate at which an inflectional form *escapes* reanalysis relative to the rate at which others do not as a function of the gang effect.

Subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ elaborates on the form of X*ó*/*æ*T-microclass members proper and satellites (see 5.2), all of which are feminine. Therefore, the combination of its dispersion with its schematicity facilitates reanalysis of plural *fætur* as a X*ó*/*æ*T-microclass form specifically. Conversely, equivalent properties as expressed in the schema [X*ur*]$_{\text{nom./acc.pl.}}$ enables it to attract and net any masculine form in plural *-ur* caught adrift in the sea of usage — all other things being equal. However, as noted in 5.3.4, this does not lead to assignment and treatment of the relevant form according to a specific feminine class. Thus, plural *vetur* and fem.pl.def. *veturnar* might just as well be substituted for *fætur* and *fæturnar* in (30), as in (31).

(31)

	[X***ur***]$_{\text{nom./acc.pl.}}$	→	[X***urnar***]$_{\text{nom./acc.pl.def.}}$
	[X*æ*T***ur***]$_{\text{nom./acc.pl.}}$	→	[X*æ*T***urnar***]$_{\text{nom./acc.pl.def.}}$
	nom./acc.pl. *stelp**ur***	~	nom./acc.pl.def. *stelpur**nar***
	nom./acc.pl. *r*æ*t**ur***	~	nom./acc.pl.def. *r*æ*t**urnar***
	↓		
	nom./acc.pl. *ve**t**ur*	→	nom./acc.pl.def. X; X = *ve**t**urnar*

Note that while plural *vetur* — also e.g. *fingur* and *eigendur* — share the ending plural *-ur* with e.g. feminine plural *stelpur*, there is little reason to assume that any of the former has been assigned to the same feminine class as *stelpa*. This function of the schema, i.e. to facilitate reanalysis on the basis of grammatical gender but not definitive class assignment, stems from its association with several classes and subclasses of Icelandic feminine nouns, in which the majority of functionally equivalent forms lack clear phonetic definition beyond plural *-ur*.

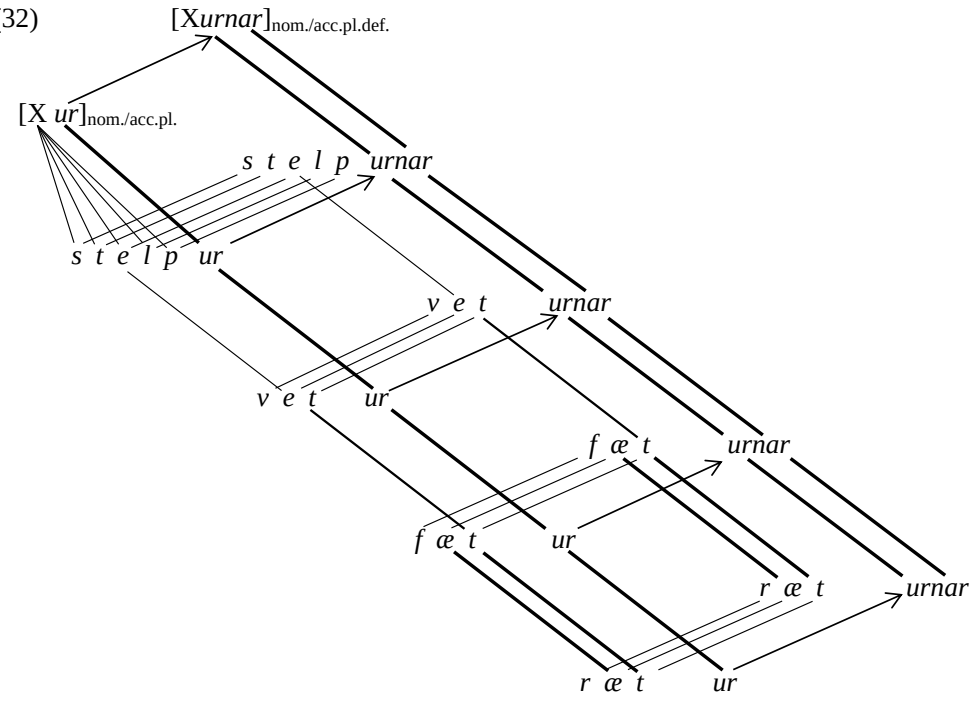
Considering now the often-observed correlation between productivity and type frequency, we might expect to associate subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ with a total lack of productivity, given the combination of its low type frequency — it represents a six-strong prototype and potentially several more with satellites of various proximity — and is minimally schematic. However, as noted by Bybee (2001:29), a pattern is capable of productivity if it is instantiated by a minimum of six items, while a minimum of two items likely suffices to facilitate the abstraction of a schema (Bybee 2010:64).[61] Therefore, by the same token, we might also expect to attribute greater productivity as a correlate of token frequency to [X*ur*]$_{\text{nom./acc.pl.}}$, relative to that of [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, on account of the former's higher type frequency and greater degree of schematicity — all other things being equal.

In light of the preceding discussion, then, the mismatch in relative token frequencies between fem.pl.def. *veturnar* and *fæturnar*, on the one hand, and their respective masculine doublet forms, on the other, appears to run counter to the view that high productivity correlates with high type frequency and high schematicity. However, we know that cross-linguistically, minimally schematic classes of medium to low type frequency display limited degrees of productivity (see the relevant discussion in 3.2.3). Therefore, Paper I argues that minimal schematicity is the very property that accounts for different degrees of productivity.

On account of the schematicity of basic level [X*ur*]$_{\text{nom./acc.pl.}}$, alignment with the schema is predicated on the ending plural *-ur* only. In other words, word-internal attributes beyond this sequence need not exhibit multiple one-to-one correspondences across paradigms for successful alignment. Conversely, alignment with [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ requires a far greater degree of one-to-one correspondence between individual segments. The nature of alignment between plural *stelpur*, *vetur*, *fætur*, *rætur*, and the schema [X*ur*]$_{\text{nom./acc.pl.}}$, as well as the schema's category validity for intra-paradigmatic alternation of the kind [X*ur*]$_{\text{nom./acc.pl.}}$ ~ [X*urnar*]$_{\text{nom./acc.pl.def.}}$, is depicted in (32) (see below on the notations employed).

[61] In support of such claims, it is noteworthy that the six-strong X*ó*/*æ*T-microclass has demonstrated some — albeit highly limited — productivity, while the semantic class that contains Ice. *kýr* 'cow' and *ær* 'ewe', i.e. two items after Ice. *sýr* 'sow' fell out of use, is completely unproductive.

(32)

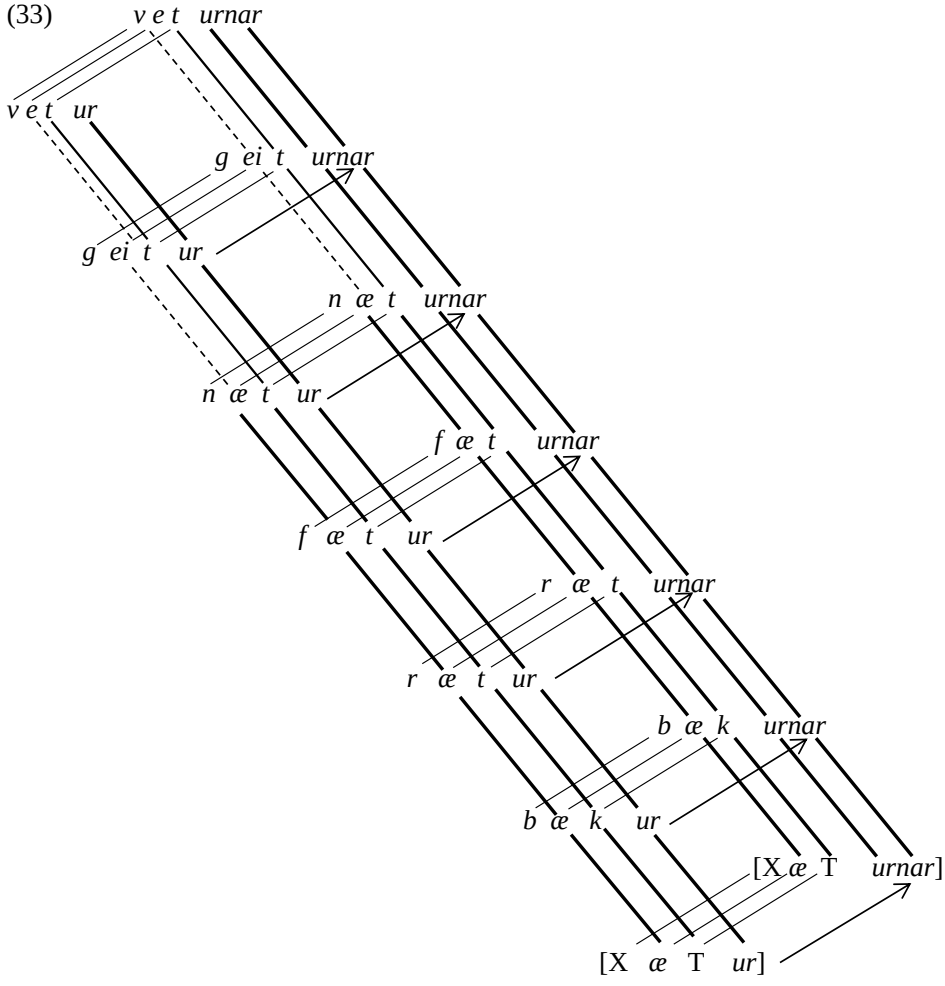


Common formal attributes that conform with schemas in terms of grammatical function are connected by thick lines, cf. connections between instances of plural *-ur*; the feminine article pl. *-nar*. Arrow-headed lines symbolise the category validity of the ending plural *-ur* for intra-paradigmatic alternation with a form in the feminine article plural *-nar*. Thin lines render arbitrary phonetic identity, which, by itself, is void of both cue validity and category validity.

Note that despite the functional role of the vowel *æ* in *fætur* and *rætur* as a marker of plural in members of the X*ó*/*æ*T-microclass, cf. sg. *fót-*, *rót-* ~ pl. *fæt-*, *ræt-*, (32) implies that this phonetic attribute is not specified at the basic level. In other words, plural *vetur* is just as likely to be reanalysed as feminine on analogy with e.g. plural *rætur* as is plural *fætur* due to formal identity between instances of plural *-ur*. Therefore, Paper I argues that at the basic level of the taxonomy posited in 5.3.2, common formal attributes beyond plural *-ur* do not impact the rate of reanalysis because all other things are equal.

To what extent, then, does disproportionate potential for structure mapping justify positing the X*ó*/*æ*T-microclass as a prototype within the feminine subtype in plural *-ur*, when the ending plural *-ur* alone suffices to facilitate reanalysis of masculine forms as feminine? To answer this question, let us now compare the implications of (32) and those inherent to (33).

(33)



On inspection of (33), the question likely arises as to why the network implies a greater degree of functional similarity between plural *nætur* and *fætur* (and, therefore, also between *fætur*, *rætur*, and *bækur*), on the one hand, than between plural *vetur*, *geitur* and *nætur*, on the other. Below, it is argued that this disparity stems from the prototype structure of the network, as the degree of schematicity decreases on convergence with the schema(s) [X*æ*T*ur*(-*nar*)]$_{\text{nom./acc.pl.(def.)}}$ as a property of family resemblance structure (see 3.3).

Thus, while common stem-final *-t* alone does not suffice to imply a form-function correspondence between plural *vetur, geitur* and *nætur*, its co-occurrence with the preceding *e* [ɛ:] in *v**e**tur*, the diphthong *ei* [**e**i:] *g**ei**tur*, the diphthong *æ* [a**i**:] in *n**æ**tur*, and plural *-ur* potentially facilitates graded membership of the microclass. In other words, the sequence *-et-* in *vetur* [v**ɛ**:tʏr] bears some similarity to the sequence *-eit-* of *geitur* [c**e**i:tʏr] due to the similarity of open-**mid front unrounded** [ɛ:] and the initial quality of the diphthong [**e**i:], i.e.

close-**mid front unrounded** [e]. Subsequently, the latter quality of the diphthong [e**i**:] shares near-high near-front unrounded [i] with the diphthong [a**i**:]. Thus, (33) allows for the possibility that [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ is schematic enough to facilitate the reanalysis and use of plural *vetur* as feminine on account of its (albeit tentative) similarity to the periphery of the X*ó*/*æ*T-microclass, via the connection of plural *geitur* to *nætur* and, thence, *fætur*, *rætur*, *bækur*, etc.

However, on account of the high degree of similarity between plural *nætur*, *fætur*, *bætur*, *rætur* as depicted in (34), Paper I argues that [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ nets the forms that align with it fully in terms of *both* form *and* function at a level of abstraction that closely reflects physical manifestation. To incorporate a maritime metaphor again: The greater the number of one-to-one form-function correspondences between a set of forms, the greater the ease with which they are *netted* by the relevant schema when caught adrift in the sea of usage. Therefore, the relatively greater number of one-to-one connections can be viewed as a metaphor for a net, the lines of which are ideal to both "catch", i.e. attract, and "release", i.e. facilitate the use of, inflectional forms as members of the X*ó*/*æ*T-microclass.

In this sense, subordinate [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, unlike its mother, basic level [X*ur*]$_{\text{nom./acc.pl.}}$, is capable of fulfilling the main function of categorisation, i.e. uncertainty reduction, at a minimal level of abstraction. That is, the one-to-one form-function correspondences in (33) that are predicated on structure mapping between instances of plural *-ur* only are just as strong as their form-function parallels in (32). However, minimally schematic [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ accurately reflects additional points of similarity between a particular set of forms in (33). In light of the above, an answer to the question posed in (27), repeated in (34) for convenience, can be given.

(34) How do we account for the frequency relation between fem.nom./acc.pl.def. *fæturnar* and *veturnar*, on the one hand, and masc.nom.pl.def. *fæturnir*, masc.acc.pl.def. *fæturna* and *veturnir*, *veturna*, on the other?

Paper I concludes that masculine plural *fætur* is relatively less likely to escape alignment with the schema [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ on both formal and functional grounds than is *vetur* or, indeed, any other masculine form in plural *-ur*. This is due to the relatively higher number of one-to-one form-function correspondences between forms that align with prototypical, functionally equivalent forms of the X*ó*/*æ*T-microclass. The greater the number of one-to-one correspondences, the greater the potential for successful alignment between functionally

equivalent forms. Therefore, the limited productivity exhibited by the low-type frequency X*ó*/*æ*T-microclass is predicated on the gang effect as a function of minimal schematicity.

5.3.6 Section summary

Subsection 5.3.1 demonstrated that frequency of use impacts categorisation via the domain-general cognitive process of statistical learning. Subsection 5.3.2 elaborated on cognitive economy as a property of functionally related linguistic categories. Subsection 5.3.3 presented instances of actual language use as evidence for hierarchical relations between the schemas that constitute linguistic categories. In subsection 5.3.4, I argued that minimal schematicity facilitates limited productivity in the absence of significant type frequency. Subsection 5.3.5 modelled the extent to which varying degrees of schematicity facilitate different rates of productivity for schemas represented at distinct levels of abstraction by notational conventions characterised as the 'net effect'.

5.4 Paper II: Deduction of plural blækur and kækur

This section elaborates on the content of Paper II. Subsection 5.4.1 presents the content of Paper II. In subsection 5.4.2, I argue for the deduction of plural *kækur* and *blækur* from singular forms of the respective paradigms as part of a two-step process. Subsection 5.4.3 details the different analogical means by which individual forms of a paradigm can be deduced, i.e. by extension and levelling. Subsection 5.4.4 argues that plural (non-humorous) *blækur* and (humorous) *kækur* are deduced by extension, negating the position that one is a better example of "real" language use than the other. Section 5.4.5 provides a summary of the current section.

5.4.1 Focus: Semantics, pragmatics, and schematicity

This and the following subsections elaborate on the content of Paper II, the focus of which is the assignment of feminine grammatical gender to forms of borrowed Ice. *blók* 'non-entity, wretch' and *kók* 'Coke™'. Inflection of both nouns occasionally follows that of the X*ó*/*æ*T-microclass. The innovations that manifest such use are feminine (humorous) plural *kækur*, beside (rare) feminine plural *kókir*,[62] and (standard) feminine plural *blækur*, beside feminine *blókir*. The paradigms of Ice. *kók* and *blók*, including doublet forms, are shown in (35), beside feminine *bók*.

[62] As noted by Svavarsdóttir (2014), masculine forms of *kók* also occurred, mainly from before and around 1960. However, these forms have since fallen out of use.

(35)			neut./fem.	fem.	fem.
	sg.	nom.	*kók*	*blók*	*bók*
		acc.	*kók*	*blók*	*bók*
		dat.	*kóki*/*kók*	*blók*	*bók*
		gen.	*kóks*	*blókar*	*bókar*
	pl.	nom./acc.	*kók*/*kækur*, *kókir*	*blækur*, *blókir*	*bækur*
		dat.	*kókum*	*blókum*	*bókum*
		gen.	*kóka*	*blóka*	*bóka*

Paper II makes clear from the outset that plural *kækur* is a humorous innovation and that some consider its occurrence not to constitute "real language use", implying that deduction of the form says little of the — albeit highly limited — productivity of the X*ó*/*æ*T-microclass (see PII:195, footnote 1). However, such reasoning fails to account for the innovative relation nom./acc.sg. *kók* ~ plural *kækur*, which is clearly modelled on e.g. *b**ók*** ~ *b**ækur***. Further, assuming that humour in some way impedes actual language use begs the question in (36).

(36) In what way is the deduction of non-humorous plural *blækur* from sg. *blók*(*ar*) a better example of "real language use" than is the deduction of plural *kækur* from sg. *kók*?

Paper II seeks to account for the occasional inflection of both Ice. *kók* and *blók* on analogy with schemas for the X*ó*/*æ*T-microclass. To this end, the objective of Paper II (see PI:196) is restated in (37), repeated from (3).

(37) To provide a usage-based cognitive account of the limited productivity of the Icelandic X*ó*/*æ*T-microclass through recourse to:

i. Semantics,
ii. Pragmatics,
iii. Schematicity,
iv. Analogy.

It is argued that deduction of both plural *kækur* and *blækur* is predicated on a high degree of phonetic similarity between both singular *kók* and *blók* and functionally equivalent forms of

the X*ó*/*æ*T-microclass. However, where treatment of individual forms of the respective paradigms differs, explanation refers to the relation between meaning and form. As information pertaining to form and function is considered to be based on stored experience with language, the mechanism posited for innovation through extension of the relevant knowledge is the domain-general cognitive process of analogy.[63]

5.4.2 Deduction as a two-step process

The current subsection argues that deduction of plural *kækur* and *blækur* from their respective intra-paradigmatic sources, i.e. singular *kók* and *blók*(*ar*), proceeds as a function of alignment with the X*ó*/*æ*T-microclass. Further, I justify positing extension as the mechanism for innovation as part of a two-step process that begins with the projection of feminine grammatical gender onto singular forms of the paradigm. Ultimately, deduction of plural *kækur* and *blækur* accords with the inflectional pattern represented by the sister schemas $[X\acute{o}T(ar)]_{sg.}$ ~ $[X\text{æ}Tur]_{nom./acc.pl.}$.

As explained in 3.2, sister schemas abstract over intra-paradigmatic relations that actually occur in the language. Further, due to the category validity inherent to *output* schemas, which activates one constituent form as implicatory of the formal attributes of another, the relevant relation is extended to new contexts, where its formal attributes come to define a functionally equivalent relation within a new paradigm, e.g. nom./acc.sg. *kók* ~ plural *kækur*, sg. *blók*(*ar*) ~ plural *blækur*, cf. *bók* ~ ***bæk*ur**. In the current context, then, the schema $[X\acute{o}T(ar)]_{sg.}$ is posited as the output schema, i.e. the source for plural targets of the kind $[X\text{æ}Tur]_{nom./acc.pl.}$.

This characterisation of the process pre-empts questions that may arise as to whether all forms of the respective paradigms are affected simultaneously with assignment of feminine grammatical gender. In other words, is the deduction of plural *kækur* and *blækur* part of a simultaneous, i.e. one-step, process or a two-step process like that argued for in Paper I (PI:9–10)? In answer, first, it is important to note that reanalysis is assumed to occur on-line, i.e. during actual usage events (e.g. Rácz et al. 2020). Secondly, it is assumed that establishment of multiple forms as constituents of a single paradigm cell is predicated on repeated reference to the relevant sister schema(s) over time, i.e. constitutes a gradual process. In terms of potential counter evidence provided by other forms of the paradigm, i.e. for wholesale, one-step gender assignment, some forms prove entirely uninformative. For

[63] For convenience, the definition of analogy is repeated here (from (6) in 3.1): [T]he cognitive process through which existing knowledge is extended to new contexts.

example, as noted in Paper I (PI:7), the endings dat.pl. *-um* and gen.pl. *-a* are all but omnipresent across the paradigms for Icelandic nouns. Further, the endings in question are common to Icelandic nouns of all three genders and, therefore, do not speak to reanalysis.[64]

Moreover, wholesale reanalysis of *kók* as feminine should yield syncretic nom./acc./dat.sg. *kók* and also gen.sg. *kókar*, an intra-paradigmatic relation of high type frequency among Icelandic strong feminine nouns, cf. e.g. nom./acc./dat.sg. *bók* ~ gen.sg. *bókar*, *mynd* ~ *myndar*. However, only one example of gen.sg. *kókar* was returned from a search of the electronic text corpus *Tímarit.is*,[65] along with one example of gen.sg.def. *kókarinnar*. Therefore, it appears that reanalysis as feminine can and does occur independently of other forms of Ice. *kók*. In this connection, it has been demonstrated that forms on opposite sides of the singular-plural divide within noun paradigms are semantically less related than forms that share a number value, i.e. either singular or plural. Such intra-paradigmatic dynamics are reflected in the relation between meaning and form, which is characterised by a tendency to align common semantics with formal similarity (see Chapter 4). Therefore, on a theoretical basis, it can be argued that the semantic relatedness of neut. dat.sg. *kók**i*** and gen.sg. *kók**s***, on the one hand, and fem. plural *kók*, on the other, to nom./acc.sg. *kók* does not proceed from an equal footing (PII:207).

Paper II argues that this mismatch in relatedness is reflected by the higher token frequency of feminine plural *kækur* relative to that of fem. nom./acc./dat.sg. *kók* and/or gen.sg. *kókar*: plural *kækur* is disproportionately more frequent than the other two forms relative to the frequency of their neuter doublets. But if the entire paradigm underwent reanalysis as part of a one-step process, we should expect feminine doublets to occur with something like proportionate frequency. Further, if semantic relatedness between functionally distinct forms is skewed, that dynamic should be reflected by a disproportionate effect of the source of change on different targets within the paradigm. As this seems to be the case with different forms of Ice. *kók*, appeal to reanalysis as a two-step process appears to be justified by statistical data.

[64] A notable exception being dative plural forms in *-m* of *some* nouns with a stem-final long vowel, cf. *skó**r*** 'shoe', dat.pl. *skó**m***, and weak feminines and neuters in gen.pl. *-na*, cf. fem. *saga* 'story, history', gen.pl. *sag**na***; neut. *auga* 'eye', gen.pl. *aug**na***.

[65] A description of the corpus taken directly from its website: "**Timarit.is** is a digital library where millions of pages in digital format are made available on the Internet. This gives access to the printed cultural heritage that is preserved in newspapers and periodicals of the Faroe Islands, Greenland, and Iceland." See the bibliography for the relevant URL.

5.4.3 Different mechanisms of analogy: Ice. kók

The current subsection details the different analogical means by which individual forms of a paradigm can be deduced, i.e. by extension and levelling. The motivation for positing different mechanisms rests on the interaction of schematicity and semantics with pragmatics, as these factors facilitate graded membership of the minimally schematic, low-type frequency X*ó*/*æ*T-microclass. To demonstrate this impact, Paper II provides examples, which are repeated and discussed below, of conventionalised phrases in which Ice. *kók* occurs as feminine in the count noun sense, a context in which this semantic property interacts with schematicity to facilitate the deduction of plural *kækur* for humorous effect.

Concerning innovative feminine plural *kækur*, it is possible to point to morphosyntactic relations that facilitate the role of nom./acc.sg. *kók* as the source of alignment with the sister schema $[X\acute{o}T]_{\text{nom./acc./dat.sg.}}$ ~ $[X\text{æ}T\textit{ur}]_{\text{nom./acc.pl.}}$. In this connection, a particular morphosyntactic context appears conducive to reanalysis of nom./acc.sg. *kók* as feminine via the pragmatic interaction of morphophonology with agreement. Thus, the following account also appeals to the use of feminine modifiers with neut. *kók* and e.g. neut. *rauðvín* 'red wine'. As noted in Paper II (PII:202), it is common idiomatic practice when ordering beverages in Icelandic to inflect modifiers of the liquid according to the grammatical gender of the container in which the former is sold. Both neut. *kók* and *rauðvín* are typically sold in a *flaska* 'bottle', while *kók* is also sold in a *dós* 'can'.[66] The words for both vessels are grammatically feminine in Icelandic and, therefore, sentences like those in (38) occur both naturally and frequently within the appropriate context (see Svavarsdóttir 2014).[67]

(38)	a.	*Ég*	*ætla*	*að*	*fá*	*eina*/ *tvær*	*kók*
		I	intend	to	get	one-fem.acc./two-fem.acc.	Coke™-acc.sg./acc.pl.
		'I'll have one/two Coke(s)™'					
	b.	*Ég*	*ætla*	*að*	*fá*	*eina*/ *tvær*	*rauðvín*
		I	intend	to	get	one-fem.acc./two-fem.acc.	redwine-acc.sg./acc.pl.
		'I'll have one/two red wine(s)'					

[66] Since the publication of Paper II, Sigurðsson et al. (2022) have published on agreement with 'concealed' or 'elided' — in my terminology 'omitted' — nouns in both what they characterise as 'restaurant talk', i.e. ordering food and beverages, and beyond.

[67] It is also common practice to order both beverages in a neut. *glas* 'glass', by use of the compounds neut. acc.sg./pl. *kókglas*/*-glös* 'glass(es) of Coke™', *rauðvínsglas*/*-glös* 'glass(es) of red wine'. In such instances, the modifying numerals always take the neuter form, cf. neut.acc. *eitt*/*tvö* 'one/two'.

Should acc.sg. *kók* be treated as feminine,[68] it might not seem unnatural on schematic grounds that language users should occasionally extend the functionally equivalent schema [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ to the relevant cell of the paradigm, given the morphosyntactic context in (38a). In other words, once acc.sg. *kók* is formally aligned with the schema [X*ó*T]$_{\text{nom./acc./dat.sg.}}$, feminine grammatical gender is projected onto the former and, thus, the latter attracts acc.sg. *kók* for the anticipated usage event. Subsequently, as functionally equivalent forms of the X*ó*/*æ*T-microclass already alternate with a plural of the type [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, cf. feminine plural *bækur*, *bætur* etc., it is a short step to deducing plural *kækur* as the target of the (now) feminine source acc.sg. *kók*, cf. (39), whether for humorous purposes or otherwise.

(39)

	[X*ó*T]$_{\text{nom./acc./dat.sg.}}$	→	[X*æ*T*ur*]$_{\text{nom./acc.pl.}}$
	acc.sg. *bók*	~	nom./acc.pl. *bækur*
	↓		
	acc.sg. *kók*	→	nom./acc.pl. X; X = *kækur*

Thus, feminine grammatical gender is projected via formal alignment of C with A, both of which align with the schema [X*ó*T]$_{\text{nom./acc./dat.sg.}}$. Subsequent to projection, the C-form also aligns functionally with the schema, which has category validity for alternation with B, i.e. *bækur*. Finally, C triggers a pattern of alternation analogous to A ~ B, hence, Paper II asserts, 'X = plural *kækur*'.

The above analysis suggests that the minimal schematicity of the X*ó*/*æ*T-microclass facilitates a significant degree of structural alignment with the sister schemas [X*ó*T]$_{\text{nom./acc./dat.sg.}}$ ~ [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ as a function of parallel connectivity. According to this line of reasoning, the same analogical process accounts for the general assignment of Ice. *blók* to the X*ó*/*æ*T-microclass. Indeed, as is argued in the next subsection, categorisation and use of *blók* as a fully-fledged X*ó*/*æ*T-microclass member serves to negate the view that deduction of plural *kækur* from singular *kók* is a substantively different process from that which yields plural *blækur* from singular *blók*(*ar*).

At this juncture, it is important to recognise the formal difference between the source of change, i.e. singular *kók*, and the target, i.e. plural *kækur*. As argued in Paper II (PII:203–204), this outcome is facilitated by a number of factors. First and foremost, from the

[68] Note that, in terms of agreement, the mass noun *kók* is otherwise inherently neuter, cf. a sentence such as *Hvar er kók**ið**?* 'Where is the Coke™?', where *kók* is clearly neuter due to the form of the definite article, nom./acc.sg. *-ið*. This is also the case when Ice. *kók* refers to cocaine.

perspective of schematicity, deduction of the innovative feminine plural form results from alignment of singular *kók* with the output schema [X*ó*T]$_{\text{nom./acc./dat.sg.}}$, as discussed above. However, when considered in the context of semantics and pragmatics, the context of use exemplified in (38a) is conducive to the expression of a semantic distinction between forms of *kók* as used in the count noun sense and, in turn, the establishment of formal differentiation of the relevant forms based on that semantic distinction. In other words, it is argued that differentiation of forms in a context such as (38a) is a function of the relation between meaning and form.

The developments just discussed are juxtaposed in Paper II with an operation imposed on the singular dative form of Ice. *kók* when used in the count noun sense. Sentences such as *vodki*/*romm í* (dat.sg.) *kók* 'vodka/rum in (idiom. *and*) Coke™' occur relatively commonly, where the preposition *í* requires a dative object. Therefore, a pertinent question given the theme of Paper II can be stated as in (40).

(40) Does dat.sg. *kók* in the sentence *vodki í kók* reflect an uninflected form of *kók* or should it be analysed as fem. dat.sg. *kók*?

Paper II (PII:205–210) argues for the latter reading through reference to schematicity and analogy based on semantic and pragmatic similarity to contexts such as (38a). Thus, Paper II contends that sentences such as *vodki*/*romm í kók* likely reflect the mass noun sense of Coke™, i.e. an indeterminate amount of the liquid used as a mixer with a measure of spirit. However, while the mass noun sense is typically expressed through use of neuter forms, a semantic bridge between such usage and that exemplified in (38a) lends itself to the analysis argued for here. In other words, the count noun sense is understood through conventionalised agreement with the feminine words *flaska* and *dós*, i.e. as portions of the liquid served in the relevant containers. This reading is obviated by the use of feminine modifiers with acc.sg., acc.pl. *kók*, as exemplified in (38a), despite the conventionalised omission of feminine *flaska* and *dós*.

Like *flaska* and *dós*, Ice. *glas* is a container that is used to serve liquid, such as a shot/measure of spirit along with a mixer. In other words, conventionalised language use like the context in (38a) is considered frequent and, therefore, entrenched enough to exert influence on related contexts, such as when ordering Coke™ as a mixer. This analysis is supported by the occurrence of PPs such as *í* ***kaldri*** *kók* 'in cold Coke™', where *kaldri* is the feminine dative singular form of *kaldur* 'cold'. In answer to the question posed in (40), then,

language use of this kind supports the argument that dat.sg. *kók* is a feminine — rather than an uninflected form. Given this interpretation, Paper II asks which analogical mechanism best characterises the deduction of fem. dat.sg. *kók* and which form of the paradigm is the most likely source of change.

Paper II (PII:206–207) argues that fem. (nom./)acc.sg. *kók*, which occurs in the influential context exemplified in (38a), is the source of the innovation on semantic grounds. In other words, when acc.sg. *kók* is used to express the mass noun as served in a feminine *flaska* or *dós*, Ice. *kók* is treated as feminine as obviated by modifiers such as fem.acc. *eina*, *tvær*. Subsequently, contexts of use in which *kók* expresses the mass noun and omitted neuter *glas*, i.e. 'a glass of Coke™', are often aligned with contexts such as that exemplified in (38a) on schematic, semantic, and pragmatic grounds.[69] It is, therefore, through this process of alignment that feminine grammatical gender is generalised to *kók* in both contexts. Further, this position opens the door for account via reference to the relation between meaning and form: both the source, i.e. fem. (nom./)acc.sg. *kók*, and target, i.e. fem. dat.sg. *kók*, express the common value *singular*.

As noted in Chapter 4, the relation between meaning and form is reflected in greater formal similarity to the extent that forms of the paradigm express a common, semantically significant value, such as a number value in nouns. Thus, in light of common semantic and pragmatic motivation for alignment of phonetic structure, the analogical mechanism for deduction of fem. dat.sg. *kók* is best characterised as levelling. In other words, as pragmatic function is aligned on the basis of common semantics, the phonetic structure of the forms involved converges as an expression of that function. Therefore, the respective analogical mechanisms via which distinct forms of the word Ice. *kók* are deduced take their cue from the interaction of schematicity, semantics, and pragmatics to express the relation between meaning and form.

5.4.4 Parallel deduction of plural blækur and kækur

As noted in Paper II (PII:195, footnote 1), some have suggested that humorous motive for the deduction of plural *kækur* belies "real" language use. However, Paper II argues that, while humour is indeed the likely motivation for use, schematicity, semantics, and pragmatics played their role in the deduction of singular *kók* through analogy with the sister schemas $[\mathrm{X}\acute{o}\mathrm{T}(ar)]_{\mathrm{sg.}} \sim [\mathrm{X}æ\mathrm{T}ur]_{\mathrm{nom./acc.pl.}}$. Plural *blækur* is indubitably deduced from singular *blók*(*ar*)

[69] See Hoffmann 2004, who argues for entrenchment of complex preposition via semantic analogy with more frequent, conventionalised, and functionally equivalent phrases.

via the very same means but the process is not considered to belie "real" usage. In light of the application of parallel processes in deduction, that of feminine plural *kækur* from singular *kók* — clearly modelled on the X*ó*/*æ*T-microclass — demonstrates that humorous motivation for the process is no impediment to "real" language use as a function of analogy.

Further, pragmatics may have played a role in the reanalysis of *blók* as feminine between the source language, i.e. English, and the receiving language, i.e. Icelandic. Indeed, English *bloke* can only refer to men (see PII:205, footnote 8). In this connection, various factors, such as phonetic structure and the applicability of a word meaning 'non-entity, wretch' to lowly people irrespective of gender identity may well have facilitated categorisation of Ice. *blók* as a feminine noun. While this constitutes conjecture due to the lack of documented change, it is noteworthy that Ice. plural *blækur* and *blókir* are both feminine, cf. in this regard that (nom./)acc.sg. *kók* also occasionally alternates with fem. plural *kækur* and (rare) *kókir*, cf. the paradigm in (35), suggesting that schematicity has facilitated parallel usage for both nouns.

For the reasons stated above, it is argued that parallel processes of analogical reasoning facilitate the deduction of both plural *kækur* and *blækur* from their respective intra-paradigmatic sources, i.e. singular *kók* and *blók*(*ar*). First, the formal similarity of *blók* to original members of the X*ó*/*æ*T-microclass likely facilitated attraction by the schema [X*ó*T(*ar*)]$_{sg.}$. Thus, the proportional equation in (41), which bears a strong phonetic resemblance to the proportion in (39), delineates extension of the X*ó*/*æ*T-microclass pattern to the paradigm of *blók*.

(41)

[X*ó*T]$_{nom./acc./dat.sg.}$	→	[XæTur]$_{nom./acc.pl.}$
acc.sg. *bók*	~	nom./acc.pl. *bækur*
↓		
acc.sg. *blók*	→	nom./acc.pl. X; X = *blækur*

Secondly, the fact that the feminine forms of numerals are used in sentences such as (38a), *despite* the fact that Ice. *kók* is chiefly neuter in the mass nouns sense, surely interacts with the noun's formal attributes to facilitate analogy. Therefore, Paper II (PII:216) argues that it is the interaction of these factors, i.e. schematicity, semantics, and pragmatics, which motivates the extension of the X*ó*/*æ*T-microclass schemas to other parts of the paradigm. Likewise, the fact that Ice. *blók* can refer to people of any gender *despite* Eng. *bloke* only

referring to men, surely interacts with the formal similarity of Ice. *blók* to functionally equivalent forms of the X*ó*/*æ*T-microclass, facilitating extension of its schemas to the paradigm. In this connection, then, the question posed in (36) is now repeated in (42).

(42) In what way is the deduction of non-humorous plural *blækur* from sg. *blók*(*ar*) a better example of "real language use" than is the deduction of plural *kækur* from sg. *kók*?

Paper II concludes that while deduction of plural *kækur* certainly is motivated by humour, the form is just as deducible from (nom./)acc.sg. *kók* as plural *blækur* is from singular *blók*(*ar*) due to the pragmatic and schematic parallels outlined above. Therefore, deduction of both plural forms is considered a mechanism of language use, itself a function of analogy. Further, this process satisfies the function of categorisation, i.e. to reduce uncertainty in the phenomena we encounter in our environment by attributing structure to it.

5.4.5 Section summary

The current section elaborated on the content of Paper II. Subsection 5.4.1 presented the content of Paper II. In subsection 5.4.2, I delineated the process by which plural *kækur* and *blækur* are each deduced as part of a two-step process. Subsection 5.4.3 argued for different analogical means for deduction of individual forms within the relevant paradigms. Subsection 5.4.4 argued that plural (non-humorous) *blækur* and (humorous) *kækur* are both deduced by extension, negating the position that one is a better example of "real" language use than the other.

5.5 Chapter summary

This chapter presented the analyses that constitute Papers I and II, which focus on the productivity of the Icelandic X*ó*/*æ*T-microclass to varying degrees. Section 5.1 laid out the content of both papers. Section 5.2 delineated the prototype structure of the Icelandic subtype in plural *-ur* as it centres on the feminine X*ó*/*æ*T-microclass. Section 5.3 fleshed out the analysis presented in Paper I, accounting for the linguistic and non-linguistic factors that facilitate the reanalysis of masculine forms in plural *-ur* as feminine at different rates of frequency. In subsection 5.3.1, I argued that frequency of use impacts categorisation via the domain-general cognitive process of statistical learning. Subsection 5.3.2 elaborated on the principle of cognitive economy as a property of functionally related linguistic categories. In subsection 5.3.3, I presented instances of actual language use as evidence in support of

hierarchical relations between constituent schemas of individual linguistic categories. Subsection 5.3.4 argued for the facilitating effect of minimal schematicity on limited productivity. In subsection 5.3.5, I modelled the extent to which varying degrees of schematicity facilitate different rates of productivity according to the 'net effect'. Subsection 5.3.6 offered a summary of Section 5.3.

Section 5.4 fleshed out the analysis presented in Paper II. Subsection 5.4.1 presented the content of Paper II. In subsection 5.4.2, I presented arguments forwarded in Paper II for the deduction of plural *kækur* and *blækur* from singular *kók* and *blók* by means of a two-step process characterised as analogical extension. Subsection 5.4.3 demonstrated that individual forms of a paradigm can be deduced by different analogical means, i.e. by extension or levelling. Subsection 5.4.4 argued that plural (non-humorous) *blækur* and (humorous) *kækur* are deduced by extension, negating the position that one is a better example of "real" language use than the other, when language use is viewed as a function of analogical reasoning. Section 5.4 was summarised in Section 5.4.5.

6 Faroese: Frequency as the determinant of levelling

The current chapter delineates the objectives, arguments, data, and conclusions presented in Paper III, which is written in Icelandic. The aim of this endeavour is to provide a sufficiently detailed overview of the article's content for those who do not have Icelandic reading proficiency. Section 6.1 discusses the content of Paper III in light of the theoretical focus. Section 6.2 fleshes out the divergent paths of levelling taken by the Faroese nouns *vøllur* 'field, grassy ledge on a rock face, (sports) pitch' and *fjørður* 'fjord, inlet/bay, sound/strait' (6.2.1), before discussing the impact of common semantics and/or frequency of use on the choice of basic form(s) for each paradigm. Subsection 6.2.3 provides a section summary. Section 6.3 establishes the basic forms for both paradigms based on the corpus data presented in Paper III. Subsection 6.3.1 posits acc.sg. *vøll* as basic according to an estimation of historical token frequencies for different forms of *vøllur*. Subsection 6.3.2 details the process of levelling in the paradigm of *fjørður*, where dat.sg. *firði* is posited as basic due to the relation *frequency* > *entrenchment* > *lexical strength*. Subsection 6.3.3 offers a summary of Section 6.3. Section 6.4 summarises the current chapter.

6.1 Paper III: Content and theoretical focus

As with Papers I and II, the analyses presented in Paper III are guided by the usage-based cognitive approach to language change. Unlike Paper I, and in common with part of the analysis in Paper II, Paper III focuses primarily on the direction of intra-paradigmatic levelling, i.e. the process through which a word acquires new inflectional forms based on existing stem variants of the paradigm. Crucially, levelling should not be viewed as the transition of one form into another. In other words, it would be a mischaracterisation of the process to say that e.g. older Far. dat.pl. *fjørðum* "changed into" younger dat.pl. *firðum*. Rather, the latter results from extension of the stem variant *firð-* from another cell of the paradigm, cf. existing dat.sg. ***firði***, nom./acc.pl. (henceforth 'plural') ***firð****ir*. This characterisation of levelling is supported by the fact that older and resultant doublets co-exist for centuries, cf. that e.g. original dat.pl. *fjørðum* and innovative dat.pl. *firðum* 'fjords' are still in use.

Moreover, Paper III deals with divergent processes of levelling in the paradigms of the Faroese masculine nouns *vøllur* and *fjørður*, both of which belonged to the class of *u*-stems

in Old West Nordic. What Papers I–III *do* have in common is that each seeks to account for the impact of token frequency, type frequency, and the dispersion of endings and/or stem variants on language change. As discussed in detail in 3.2.2, frequency of use impacts the domain-general process of entrenchment, which, in turn, is the determinant of lexical strength. Thus, by building on the testimony of levelling in other languages, Paper III seeks to determine the **basic forms** of paradigms, i.e. those forms upon which new ones are based.

Given the causal relation *frequency* > *entrenchment* > *lexical strength*, Paper III seeks to answer the question already posed in (4); now repeated in (43).

(43) Does frequency determine the direction of levelling?

As noted in Chapter 4, in order to answer this question, Paper III utilises modern-day token frequencies for different inflectional forms of the relevant nouns, based on results returned from the electronic corpus for Modern Faroese, *Teldutøka tekstasavn Føroyamálsdeildarinnar* (*TTF*) and, in order to establish historical token frequencies for different forms of Far. *vøllur*, a comparison with data returned from the electronic corpus for Icelandic, *Íslenskt textasafn* (*ÍT*).

In this connection, Paper III asserts that use of Far. *vøllur* for topographical referents, i.e. in the senses ‘field, grassy ledge on a rock face, (sports) pitch’, has likely always been rare due to the topography of the Faroe Islands (see PIII:62 and sources cited there). Further, it is demonstrated that Far. *vøllur* is used far more frequently in the modern language in reference to sports pitches than in any other sense at any period before the 20th century. Moreover, use of the word as a shortened form of *flogvøllur* ‘airport’ is an obvious nuance of the 20th century. In light of this, Paper III seeks to estimate the historical token frequencies for individual forms of *vøllur* through comparison with equivalent instances of use in both Modern Faroese and Old Icelandic written sources, where the latter is taken as representative of Old West Nordic more generally (see Chapter 2 on the relative scarcity of written sources for Faroese). This approach also considers real-world conditions such as topography, weather, and luminescence, as such factors would have impacted frequency of use.

Conversely, it is demonstrated that Far. *fjørður* has likely always referred to the same phenomena since before the time of settlement, i.e. ‘fjord, inlet/bay, sound/strait’. This noun is also common in place names, where it refers to one of these natural phenomena. Thus, comparison of the different directions of levelling taken by Far. *vøllur* and *fjørður* prove highly informative with regard to the validity of theoretical approaches that seek to account

for the choice of basic form(s): both *vøllur* and *fjørður* have topographical referents and occur in place names (see PIII:54).

Therefore, development of the paradigms in question provide opportunity to evaluate the impact of common semantics on the direction of levelling; a tenet of the approach to levelling associated with so-called **markedness**, which views semantically "natural", "neutral", and/or "default" inflectional categories as diagnostic of basic status (Greenberg 1966). Given this view, we might expect nouns that share a semantic domain to undergo similar changes with regard to the direction of levelling. However, as demonstrated time and again, actual language change is often at odds with the approach from markedness, which, it is argued, is ultimately characterised by circularity (see PIII:58 and sources cited there).

In light of this drawback to markedness, I chose to test the impact of frequency because it is a measurable property of usage, while semantics is not (following Bybee 2015:102; Haspelmath 2006:27, 58; Sims-Williams 2022:571). In Paper III, the term frequency (Ice. *tíðni*) is first and foremost used in reference to token frequency (Ice. *staktíðni*), which is considered to determine the initial choice of basic form as a function of the causal relation *frequency* > *entrenchment* (Ice. *rótfesta*) > *lexical strength* (Ice. *minnisstyrkur*). Further, it is argued that both the type frequency (Ice. *mynsturstíðni*) of patterns and the dispersion (Ice. *dreifitíðni*) of stem variants within the paradigm assists in completing the process of levelling as a function of the cross-linguistic tendency to establish a relation between meaning and form (Ice. *samband forms og merkingar*). Paper III concludes that the divergent processes of levelling discussed stem from the interaction of token frequency, type frequency, and (intra-paradigmatic) dispersion.

6.2 The direction of levelling

This section delineates the divergent paths of levelling taken by the Faroese nouns *vøllur* and *fjørður* (6.2.1), before discussing the impact of inhabiting a common semantic domain, i.e. having topographical references and occurrence in place names, and/or frequency of use on the choice of basic form(s) for each paradigm. Subsection 6.2.3 provides a section summary.

6.2.1 Levelling in Far. vøllur and fjørður

In the current subsection, I delineate the divergent paths of levelling taken by Far. *vøllur* and *fjørður*. Crucially, both nouns belonged to the class of *u*-stems in Old West Nordic, meaning that they showed the same endings and exhibited functionally parallel patterns of alternation (see below). The paradigms of Old West Nordic *u*-stems have undergone extensive

analogical levelling, eliminating all but one older stem variant in the vast majority of cases. However, as discussed in Paper III (PIII:55–56), the direction of levelling in the paradigms of Far. *vøllur* and *fjørður* (< OWN *vǫllr* and *fjǫrðr*) demonstrates that the process was not uniform across-the-board. In order to gain perspective in the historical context, the paradigms of OWN *vøllur* and *fjørður* are shown in (44).

(44)

sg.	nom.	*vǫllr*	*fjǫrðr*
	acc.	*vǫll*	*fjǫrð*
	dat.	*velli*	*firði*
	gen.	*vallar*	*fjarðar*
pl.	nom.	*vellir*	*firðir*
	acc.	*vǫllu*	*fjǫrðu*
	dat.	*vǫllum*	*fjǫrðum*
	gen.	*valla*	*fjarða*

For the purpose of comparison, the paradigms of Far. *vøllur* and *fjørður* are given in (45), where the stems of forms that arose through levelling are in **bold**.[70]

(45)

sg.	nom.	*vøllur*	*fjørður*
	acc.	*vøll*	*fjørð*
	dat.	***vøll****i*	*firði*/***fjørð****i*
pl.	nom./acc.	***vøll****ir*/***vall****ir*	*firðir*/***fjørð****ir*/***fjarð****ir*
	dat.	*vøllum*/***vall****um*	*fjørðum*/***firð****um*/***fjarð****um*

As obviated by its paradigm in (45), Far. *vøllur* has followed the general path of *u*-stems, i.e. levelling in favour of the stem variant found in the singular nominative and accusative and those of the plural accusative and dative, i.e. *vǫll-* (> Far. *vøll-*).[71] The variant *vell-* has been

[70] Although separate genitive forms exist in Faroese, the genitive case is no longer actively assigned, its disappearance having yielded a three-case distinction in the modern language, i.e. nouns are inflected for nominative, accusative, and dative case (Þráinsson et al. 2012:62).

[71] Extension of the stem variant *vall-* appears to be modelled on the pattern of alternation exhibited by e.g. sg. *høll* ~ nom./acc.pl. *hallir* ~ dat.pl. *høllum*/*hallum* ‘palace(s); centre(s), hall(s)’. In other words, it is probably the

completely eradicated from the paradigm. Conversely, Far. *fjørður* has retained all of the stem variants it inherited from Old West Nordic and has acquired several new forms with the variants *fjørð-*, *fjarð-* and *firð-*. It is this disparity between the two paradigms that motivates the question in (4)/(43) and now repeated in (46) for convenience.

(46) Does frequency determine the direction of levelling?

The data, their interpretation, and attendant arguments laid out in the following subsections will facilitate the answer to this question.

6.2.2 The choice of basic form(s)

In this subsection, I discuss the impact on nouns of sharing a semantic domain and of frequency of use on the choice of basic form(s) for the paradigms of Far. *vøllur* and *fjørður*. Focus on the development of these nouns was not a random choice. Rather, as noted in 6.1 and 6.2.1 the nouns both have topographical referents and appear in place names. Further, they formerly exhibited parallel inflectional attributes, both with regard to inflectional endings and alternation between stem variants. First, both nouns are descended from Old West Nordic *u*-stems, i.e. *vǫllr* and *fjǫrðr*, respectively. This means that they showed the same inflectional endings and exhibited functionally parallel patterns of alternation at an earlier stage, cf. the Old West Nordic paradigms in (44).

Secondly, Far. *vøllur* and *fjørður* have topographical referents that occur both as features of the landscape and in place names, e.g. *Harða**vøllur***, *Tórs**vøllur***; *Fugla**fjørður***, *Skopunar**fjørður*** (all place names). This point of commonality is especially significant in light of the theoretical approach to levelling that is associated with so-called markedness. The approach from markedness mainly attributes the selection of basic forms to the influence that semantically "natural", "neutral", and/or "default" inflectional values supposedly have on use, as briefly noted in 6.1. In the context of nominals, proponents of markedness consider the nominative case to possess all three qualities and, therefore, to be **unmarked** compared with other case forms, which are considered **marked** (e.g. Greenberg 1966).

By the same token, singular number is generally considered unmarked against the dual and plural. In other words, nominative and singular represent the supposedly more neutral case and number values, while other values within each category convey extra information.

result of extension rather than levelling and, therefore, falls out of the scope of the current focus (see PIII:56, footnote 5).

In metaphorical terms, this extra information often literally "marks" the forms in questions, then requiring additional phonetic material in order to express extra semantic information (see below). According to the approach from markedness, then, the nominative singular form of a "regularly inflected" noun should logically serve as basic, i.e. new forms within the paradigm should be based on the form of the nominative singular (Petersen 2009:89 states specifically that the nominative is the default case for Faroese).

However, actual language change demonstrates that levelling does not always proceed from those forms of the paradigm that proponents of markedness consider unmarked. Indeed, while some might attribute this status to nom.sg. *vøllur* in light of levelling within its paradigm, the course of change in Far. *fjørður* clearly runs counter to such a view in some respects, cf. (44) and (45). Indeed, as argued in Paper III (PIII:77), innovative dat.pl. *firðum* suggests that nominative and singular *fjørður* did *not* serve as basic with regard to levelling. So, how have proponents of markedness sought to resolve such exceptions to the general applicability of the general theory?

In order to impose systematicity on the obvious shortcomings of the approach, scholars attribute the quality of **local markedness** to paradigms where levelling runs counter to the notion of **general markedness**, discussed above. For example, Mańczak (1958) noticed that levelling in the paradigms of place names tended to favour locative forms as basic. In Faroese, the most obvious equivalent of the locative case in e.g. Russian is the dative. Indeed, objects of the prepositions Far. *í* 'in' and *á* 'on' occur in the dative to indicate location, e.g. *í* (dat.) *Fuglafirði* 'in Fuglafjørður'. Bernharðsson (2004:25–26) notes that the dative case was especially common in this function in Old Icelandic, where the preposition *að* with dative object frequently denoted location also, e.g. *lét kalla* ***að*** (dat.) *Ökrum* 'named it Akrar'. As this is the case in Old Icelandic, the same dynamic likely also applied in the contemporaneous Old West Nordic dialect spoken in the Faroe Islands (see Chapter 2).

Beyond the dative case, use of the accusative in Faroese renders a relation of increasing proximity to a location. In other words, the Faroese accusative is not prototypically locative but indicates movement towards a place (see Þráinsson et al. 2012:164). Its use in this function is also associated with the prepositions *í* and *á* but also with Far. *til* 'to', e.g. *Eg fari oman* ***í*** (acc.) *býin* 'I'm going downtown', *Vit fara* ***til*** (acc.) *Fuglafjørð* 'We are going to Fuglafjørður'.[72] Given the fact that both the dative and accusative are associated with

[72] In Old West Nordic, the proposition *til* governed the genitive case, as its Icelandic reflex does to this day. However, the accusative now frequently marks the object of *til* in Faroese (see footnote 70).

location in Faroese — albeit to different degrees — Paper III contends that we might have expected Far. *vøllur* and *fjørður* to show more uniform development with regard to levelling in light of the approach from (local) markedness. So, why is this not the case?

In advance of the explanation, it should be noted that Haspelmath (2006) considers the term *markedness* inherently vague, attributing this shortcoming to the generality of the term as applied in the linguistic literature. Thus, he demonstrates, markedness can be defined in at least twelve different ways, so that it means different things to different linguists, depending on the material chosen for analysis and the theoretical approach applied. Further, it appears from Haspelmath's survey of the term's application in linguistic inquiry that some scholars are not aware of its different uses (2006:27). In other words, "markedness" is reducible to numerous practical applications in the linguistic literature and, as a result, serves a highly general function, rather than constituting a well-defined term.

In an attempt to resolve this issue, Haspelmath (2006) notes that, in most studies where markedness is invoked, the term is reducible to frequency effects in language use and change. In this connection, it should be noted immediately that frequency has itself been considered an indicator of the (un)marked status of forms within paradigms (e.g. Greenberg 1966). However, as pointed out most recently by Sims-Williams (2022:571) and in Paper III (PIII:58–59), invocation of frequency as a diagnostic criterion for markedness renders application of the term subject to circularity.

Consider the following axiom: Inflectional values considered semantically "natural", "neutral", and/or to be "default" relative to opposing categories are unmarked. Further: Unmarked forms tend to be more frequent than marked forms. However, a logical fallacy inherent to such an approach is obvious if relatively higher frequency is subsequently taken as diagnostic of semantically more natural, neutral, and/or default status within the paradigm. In other words, the approach from markedness proves circular: Higher frequency as a property of use provides evidence of unmarked status > if a form is unmarked it will likely be more frequent than other forms of the paradigm > higher frequency as a property of use provides evidence for unmarked status. And so the cycle continues.

Paper III espouses a more falsifiable approach to determining the basic forms of paradigms. By accounting for the process of levelling in terms of the impact of frequency on the cognitive organisation of language, linguists are offered a means of disproving their hypotheses: frequency is a measurable property of language use, while meaning is not. Further, any account that seeks to attribute the status of basic form to semantics must do so on a theory-internal basis. In other words, the theory that a set of inflectional values

determines the general direction of levelling (general markedness) requires that a given word convey a meaning that permits those categories to determine the direction of levelling (local markedness). Conversely, the usage-based cognitive approach to language finds support from beyond the linguistic system, in properties such as frequency of occurrence, schematicity, and context of use, all of which are demonstrable determinants and yields of statistical learning, schematisation, and entrenchment.

Support for this view comes from various studies on the determinants of levelling. For example, demonstrating that the direction of levelling is independent of the semantic properties of individual words, Tiersma (1982) noted that basic forms for Frisian nouns were selected on the basis of their individual token frequencies. Levelling of this nature has also occurred in Icelandic, where e.g. the stem OIce. *gæs-* 'geese' from the more frequent plural has replaced OIce. sg. *gás-*, now sg. *gæs-* (see Bernharðsson 2006). The Old West Nordic vocalic alternation is preserved in the paradigm of the Faroese cognate, thus Far. sg. *gás* ~ nom./acc.pl. *gæs* ~ dat.pl. *gásum*. It is important to note here that statistical learning is a tacit process, meaning that the gradual process of entrenchment determines the ease with which a given stem variant is retrieved from memory, i.e. its emergent lexical strength (see 3.2.2). Therefore, the direction of levelling is determined by the accessibility of competing linguistic structures and the different patterns that can be applied in order to express a given concept (see discussion in PIII:75). In other words, language users are not necessarily actively aware that they are choosing the most frequent form of the paradigm as basic. Rather, the choice is informed by tacitly accrued, entrenched information pertaining to frequency.

6.2.3 Section summary

In this section, I discussed the nature of levelling and its determinants according to opposing theories. In 6.2.1, I delineated the different processes of levelling taken by Far. *vøllur* and *fjørður*, demonstrating that the former followed the majority of *u*-stems in Faroese language history, while the inflection of *fjørður* has become more complex. This disparity was considered interesting in light of semantic properties common to the nouns in question: both have topographical referents and occur in place names. In 6.2.2, I argued against semantics as a basis for positing basic forms due to the vagueness of the term "markedness". It was concluded that frequency of use provided a measurable and, therefore, falsifiable means of testing the impact of language use on lexical strength and is, therefore, a better-suited gauge for establishing the basic forms of paradigms.

6.3 Determining the basic forms of Far. vøllur and fjørður

The current section accounts for the direction of levelling based on frequency data presented in Paper III and in light of the three distinct parameters for frequency counts acknowledged in 3.2.1, i.e. token frequency, type frequency, and dispersion. In 6.3.1, I posit acc.sg. *vøll* as basic on the strength of the estimation of historical token frequencies for different forms of the word. Further, I argue that dispersion of the stem variant *vøll-* in the singular portion of the paradigm ultimately determined the direction of levelling. Subsection 6.3.2 details the process of levelling in the paradigm of Far. *fjørður*. There, dat.sg. *firði* is posited as basic. Further, levelling in favour of the dative singular is argued to have motivated a subsequent attempt to establish a relation between meaning and form that distinguished singular from plural, a process that takes its cue from dispersion. Subsection 6.3.3 offers a summary of the current section.

6.3.1 Far. vøllur

In the course of Faroese language history, the stem variant *vøll-* (< OWN *vǫll-*) has been generalised to all cells of the paradigm for Far. *vøllur*, cf. (44) and (45). As noted in 6.2.1, the development of Far. *vøllur* can be considered representative of the general path of levelling in Old West Nordic *u*-stems. This process is typically characterised by extension of the stem variant found in the singular nominative, accusative, and plural accusative and dative to all cells of the paradigm (see PIII:56 for a list of examples).

As argued in Paper III (PIII:72), such developments reflect the high type frequency of phonetic uniformity in the stems of individual Faroese masculine nouns: the vast majority contains the same stem variant in all cells of the paradigm. Further, *u*-stems shared the inflectional ending nom.pl. *-ir* with the larger class of masculine *i*-stems. Moreover, formal and functional similarity between the two classes is to some degree increased via one-to-one correspondence based on the ending gen.sg. *-ar*, which all *u*-stems shared with a small number of nouns from other masculine classes; mainly *i*-stems. Therefore, the general direction of levelling among *u*-stems in Faroese can be accounted for via reference to the interaction of high type frequency of stem uniformity in masculine paradigms and points of formal and functional convergence with a larger masculine class.

However, reference to type frequency does not suffice in accounting for the divergent course of levelling apparent from the paradigm of Far. *fjørður*. This fact is taken to suggest that type frequency alone might not be the only factor that contributes to productivity, suggesting that Yang's *Tolerance Principle* is too rigid (see footnote 16). In order to

demonstrate that other parameters of frequency are at play in the developments under discussion, Paper III posits a chronologically ordered course of levelling for Far. *vøllur* through reference to token frequencies for different forms of the word, which are based on an estimate for earlier phases of the language via comparison with contemporaneous frequencies for forms of the Icelandic cognate *-vǫllr*/*völlr*/*völlur* (PIII:68). Further, the account also avails itself of the dispersion of individual stem variants, which is said to constrain distribution across the singular-plural divide as a function of the relation between meaning and form (Chapter 4).

Based on the estimated token frequencies reported in Paper III and that are considered to reflect usage patterns before the 20th century, acc.sg. *vøll* had the highest token frequency, equivalent to 34 instances, compared with dat.sg. *velli*/*vølli*, with the equivalent of ten occurrences (see PIII:68, Table 4). Given that token frequencies for individual forms of Far. *vøllur* were highly likely to have been rare in centuries past, Paper III (PIII:75) asks whether the difference between 34 instances and 10 suffices to determine the choice of basic form. The answer to this question references the intra-paradigmatic dispersion of different stem variants: when the singular portion of the paradigm is considered as a whole, it becomes clear that older dat.sg. *velli* had not only to compete with the equivalent of 34 occurrences of acc.sg. *vøll* but also ten instances of nom.sg. *vøllur*. Both of the latter contained the stem variant *vøll-*.

Thus, in light of the view that frequency impacts the direction of levelling as a function of statistical learning, Paper III argues that the combined token frequencies of acc.sg. ***vøll*** (34 instances) and nom.sg. ***vøll****ur* (10 instances) interacted with the dispersion of the stem variant *vøll-* over two of the three cells of the singular.[73] The combined token frequency of the forms with the stem variant *vøll-* accounts for 81.48% of the token frequency for all three singular forms. Further, the same variant enjoys a 66.66% rate of dispersion in the singular portion of the paradigm, meaning that language users are more likely to associate the form *vøll-* than *vell-* with the meaning *singular*. Therefore, given the impact of frequency as a determinant of both entrenchment and lexical strength, it is argued that the stem variant *vøll-* will have been more easily sought from memory to serve this semantic function.

[73] Of course, levelling may have commenced before the genitive was lost as an active case in Faroese, meaning that the singular portion of the paradigm would then have contained four cells with three distinct stem variants, i.e. nom./acc.sg. *vøll-*, dat.sg. *vell-*, and gen.sg. *vall-*. However, the stem variant *vall-* now only occurs in plural forms of the paradigm and is the likely result of extension based on e.g. sg. *høll* ~ pl. *hallir*, as noted in 6.2.1, cf. the innovative alternation Far. sg. *bjørn* ~ pl. *bjarnir*, OWN *bjorn* ~ *birnir*.

Paper III also asks why the stem variant *vell-* was not subsequently generalised to the plural portion of the paradigm, given that it expressed the function nominative plural and possibly accusative plural too, had the latter been levelled to *velli*(*r*) before the stem variant *vøll-* started to spread. Such a dynamic would have yielded a relation between meaning and form, by which *vøll-* equated with *singular* and *vell-* with *plural*. However, as Paper III (PIII:76) notes, the individual token frequencies for plural forms of Far. *vøllur* have likely always been relatively low (cf. PIII:63, Table 1 and PIII:68, Table 4). Therefore, given the generally low frequency of all forms of the noun, the variant *vell-* is unlikely to have been entrenched enough to resist levelling to *vøll-*. Indeed, intra-paradigmatic pressure from both the singular portion of the paradigm and dat.pl. *vøllum*, which appears to have been more frequent than nom.pl. *vellir* and acc.pl. *velli*(*r*), likely facilitated levelling in favour of the stem variant *vøll-*.

It should, of course, be acknowledged that it is advisable to err on the side of caution when making claims about the historical frequency of inflectional forms and the chronology of change in the absence of primary written sources. However, Paper III argues that we must make use of the data available to us rather than foregoing an attempt at explanation. On the basis of the method for recalculation delineated in Chapter 4, the hypothesis that frequency should impact the direction of levelling is testable. In other words, the available data is not interpreted in order to fit the theoretical approach but, rather, the approach has been subsequently applied to the estimation of historical frequency based on the recalculation. The method for recalculation was motivated by a preponderance of use as affected by real world conditions and was based on a comparison of relevant — albeit secondary — sources. As also noted in paper III (PIII:66, footnote 17), the study is not the first to rely on secondary sources as a basis for estimating historical frequencies.

6.3.2 Far. fjørður

As discussed in 6.2.1, the development of Far. *fjørður* with regard to levelling is atypical of the process among *u*-stems in Faroese language history. Indeed, while almost all nouns of this class have generalised the stem variant descended from the singular nominative and accusative and plural accusative and dative in Old West Nordic, the paradigm for *fjørður* contains innovative forms with stem variants from other cells of the older paradigm, cf. (44) and (45).

As argued in 3.2.2, token and type frequency are considered the explicit mechanisms of entrenchment, which determines lexical strength. It is important to note in this connection

that while the overall course of levelling in the paradigm of Far. *fjørður* does not speak to the influence of any particular pattern of inflection, the process as it occurred within the singular and plural portions of the paradigm, respectively, does speak to such influence. Thus, the relation nom.sg. ***fjø**rður* ~ acc.sg. ***fjø**rð* ~ (innovative) dat.sg. ***fjør**ði* aligns with functionally equivalent relations in larger masculine classes, cf. *a*-st. *h**e**stur ~ h**e**st ~ h**e**sti*, *i*-st. *g**e**stur ~ g**e**st ~ g**e**sti*, where the stem is identical in all forms. Likewise, the relation nom./acc.pl. ***fi**rðir*, (innovative) dat.pl. ***fi**rðum* accords with e.g. *g**e**stir*, *g**e**stum* for the same reason. However, it is clear from data presented in Paper III (PIII:69, Table 5) that disparities in token frequency between individual forms of *fjørður* are a property of its use. It is, therefore, fair to assume a role for the impact of token frequency in the direction of levelling and investigate the hypothesis that this frequency parameter served as the initial catalyst for change.

As noted in Paper III (PIII:80), an obvious position from which to mount that investigation is to view highly frequent dat.sg. *firði* as basic in the paradigm. In this connection, it is important to note that the only other cell to which the stem variant *firð-* has been extended is that of the dative plural, cf. younger ***firð**um*, which now exists alongside older *fjørðum*. It is, of course, also possible that the stem of this younger form is derived from plural ***firð**ir*. However, Paper (PIII:77–78) presents two arguments against this proposition. First, motivation to level the plural in this way would be a function only of the dispersion of the relevant stem variant and not the impact of its frequency on entrenchment and lexical strength: dat.pl. *fjørðum* has higher token frequency than nom./acc.pl. *firðir* and, therefore, the former should be better entrenched and easier to access than the latter (see PIII:69, Table 5). In other words, while dispersion of the variant *firð-* across the plural portion of the paradigm may well have facilitated a relation between meaning and form, there is little to suggest that dispersion contributes to entrenchment and, therefore, also lexical strength at the cost of token frequency (see discussion in 3.2.1).

Secondly, the dative forms of both numbers occur in a common pragmatic context, i.e. as the objects of the prepositions *í*/*á*/*úr*, cf. the schema [*í*/*á*/*úr* FJØRÐUR$_{dat.}$], e.g. *í Fuglafirði, úr Kollafirði*. Further, the schema just posited is likely deeply entrenched due to the frequency that it is invoked on account of language users' perception of real-world conditions. In other words, *Fuglafjørður*, *Kollafjørður*, and most other place names ending in Far. *-fjørður* refer to both fjords and towns, where people live and locate events in reported speech. If it is correct that frequency of use is the mechanism of entrenchment, then it is logical to assume that the impact of frequency on levelling is a function of the relative ease

with which a stem form is accessible for use, i.e. that form's lexical strength, within a particular context.

In other words, invocation of the schema [*í*/*á*/*úr* FJØRÐUR$_{dat.}$] is a highly practiced behaviour, which, due to the process of automatisation, becomes more effortless with time (see 3.2.1). One recognised function of automatisation is **autonomy**, by which a form or lexical item acquires a high degree of independence from the functional constraints of its paradigm or common context of use and, in emergent fashion, becomes less compositional, i.e. analysable in terms of its constituent morphological elements (e.g. Bybee 2010:48). Therefore, it is likely that due to the high frequency with which the schema [*í*/*á*/*úr* FJØRÐUR$_{dat.}$] contains the form dat.sg. *firði*, the stem variant *firð-* will have become deeply entrenched in this context, irrespective of its function in relation to other forms of the paradigm.

Paper III (PIII:75) also argues that, as a function of rich memory for language, individual form-function units, such as stem variants and inflectional endings, are stored in memory both separately, cf. e.g. [*firð-*]$_{dat.sg., nom./acc./dat.pl.}$ and [*-ir*]$_{nom./acc.pl.}$, [*-um*]$_{dat.pl.}$, and in broader contexts as constituent parts of the word, cf. dat.sg. *firði*, nom./acc.pl. *firðir*, or dat.pl. *firðum*. Indeed, were the individual building blocks of word forms not stored in memory, it should not be possible to attach the ending dat.pl. *-um* to the stem variant *firð-*, yielding innovative dat.pl. *firðum*. Further, were the process not a function of the impact of frequency on use, we should expect to witness equivalent processes for all forms of Far. *fjørður*, as none would be subject to change as a property of stored experience but, rather, on-line language processing alone. However, the development of the relevant paradigm suggests that forms such as nom.sg. **firður* and acc.sg. **firð* do not occur, a property of usage attributable to the fact that there is no crossover between nom.sg. *fjørður* and acc.sg. *fjørð*, on the one hand, and dat.sg. *firði*, on the other, within contexts such as [*í*/*á*/*úr* FJØRÐUR$_{dat.}$].

Given the above discussion, Paper III concludes that younger dat.pl. *firðum* is the first form of Far. *fjørður* to result from the process of levelling within the paradigm (PIII:77–78). The reasoning for this position is based on entrenchment, in light of the fact that the human cognitive faculties have been uniform across the species for millennia. In other words: the domain-general mechanisms of language use at play today are the same that have propelled language change in all language communities at all times (Bybee 2003). Thus, it is argued that Far. *firði* has always been the most frequent member of the paradigm across a period well beyond the living memories of contemporaneous speakers at any time since the Faroe Islands were settled. Thus, in light of the causal relation *frequency > entrenchment > lexical*

strength, it is also reasonable to assume that the same form has always been more entrenched than other forms of Far. *fjørður* at all stages in Faroese language history.

It is argued that token frequency as a mechanism of levelling facilitated extension of the stem variant *firð-* from dat.sg. ***firði*** to the dative plural form — especially within the context [*í*/*á*/*úr* FJØRÐUR$_{dat.}$] — potentially even relatively early in Faroese language history. In light of this position, extension of the stem variant *fjørð-* to the dative singular cell, cf. younger dat.sg. ***fjørði***, is considered to result from a reactionary operation: a function of the relation between meaning and form, rather than the lexical strength associated with any of the forms that originally contained that variant (see PIII:80). Once dat.pl. *firðum* comes into use, it soon gains in token frequency due to the degree to which its stem is entrenched in the context [*í*/*á*/*úr* FJØRÐUR$_{dat.}$], facilitating association of the stem variant *firð-* with the meaning *plural*, cf. nom./acc.pl. ***firð****ir* and (innovative) dat.pl. ***firð****um*. This process yields an uneven relationship between meaning and form because the stem variant *firð-* still exists in the singular portion of the paradigm, cf. dat.sg. ***firð****i*, opposing both nom.sg. *fjørður* and acc.pl. *fjørð*.

Paper III concludes, therefore, that users of Faroese subsequently attempted to establish a relation by which the stem variant *fjørð-* was equated with the meaning *singular*, the logical solution to which was to derive younger dat.sg. ***fjørði*** by borrowing the stem variant from nom.sg. ***fjørð****ur* and acc.sg. ***fjørð***. However, it is argued that — despite language user's best efforts — the token frequency of original dat.sg. *firði* was (and still is) so much higher than that of the other singular forms combined that dispersion of sg. *fjørð-* does not suffice to usurp the function of dat.sg. *firði*. This position is supported by the fact that the former has not taken the latter over in terms of token frequency, while younger dat.pl. *firðum* is far more frequent than older dat.pl. *fjørðum*, probably due to the high token frequency of dat.sg. ***firð****i* and the resulting entrenchment of the stem variant *firð-* in the context [*í*/*á*/*úr* FJØRÐUR$_{dat.}$].

6.3.3 Section summary

The current section provided an account of the process of levelling in the paradigms of Far. *vøllur* and *fjørður* as depicted in Paper III, which sought to determine the basic forms of each. Subsection 6.3.1 argued that acc.sg. *vøll* served as the basic form of Far. *vøllur* due to its historically higher token frequency. Further, it was argued that dispersion of the stem variant *vøll-* across two of the three cells of the singular portion of the paradigm served to facilitate levelling in favour of the stem variant *vøll-* generally. In subsection 6.3.2, I detailed levelling in Far. *fjørður*. There, dat.sg. *firði* was posited as basic, while extension of its stem

variant was argued to have been countered by an attempt to establish a relation between meaning and form that distinguished singular from plural. The process that derived innovative dat.sg. *fjørði* is regarded to have taken its cue from the dispersion of different stem variants within each number.

6.4 Chapter summary

The current chapter delineated the contents of Paper III. Section 6.1 fleshed out the contents of the article in light of the theoretical focus. Section 6.2 delineated the divergent paths of levelling taken by the Faroese nouns *vøllur* and *fjørður* (6.2.1), subsequently discussing the impact of common semantics and/or frequency of use on the choice of basic form(s) (6.2.2). Subsection 6.2.3 provided a summary of Section 6.2. Section 6.3 posited the basic forms for both paradigms based on frequency counts from a text corpus for Faroese. Subsection 6.3.1 argued that acc.sg. *vøll* is basic to the paradigm of Far. *vøllur* according to an estimation of historical token frequencies for different forms of the word. Subsection 6.3.2 detailed the process of levelling in the paradigm of Far. *fjørður*, where dat.sg. *firði* was posited as basic. Subsection 6.3.3 offered a summary of Section 6.3.

7 Conclusion

The current thesis sought to investigate and demonstrate the impact of domain-general cognition as a determinant of language change, itself a function of language use. To this end, changes in the inflectional systems of Icelandic and Faroese have been in focus, with the yield of these endeavours being presented in Papers I–III (see 1.1 for bibliographical information). As noted in 3.1, research conducted over the last decades into the grammatical system of Icelandic has primarily been guided by the tenets of generative and dual-processing models. Conversely, application of the usage-based cognitive approach has been mostly implicit, i.e. without statement of specific tenets of the relevant model. In this context, it is hoped that the current thesis motivates further research into the respective inflectional systems of Icelandic and Faroese from the usage-based cognitive perspective.

As noted throughout, the current thesis argues for the impact of domain-general cognitive processes such as statistical learning, entrenchment, categorisation, and schematisation as facilitators of both language use and language change. Further, in order to show that these processes apply domain-generally, it was paramount that the approach demonstrate the impact of non-linguistic factors such as frequency and schematicity in language change. Moreover, demonstration of this impact is viewed as evidence for rich memory for language, thus calling the validity of more parsimonious theoretical models, such as dual-processing, into question. Indeed, as noted in 3.2.1, were a robust capacity for the storage of linguistic experience not a property of memory, we should expect to witness far fewer — if any — observable frequency effects in language change.

In the context of the above, the current thesis set out to satisfy the general objectives stated in (1a–c), which, for convenience, are repeated below in (47a–c).

(47) a. To determine which cognitive factors have impacted some details of the structure of the inflectional systems of Icelandic and Faroese.

b. To ascertain whether these cognitive factors apply only to language or are domain-general in nature.

c. To demonstrate how the impact of these cognitive factors are evident through language change as a function of language use.

In light of the aims reiterated above, the analyses presented in Papers I–III are considered to satisfy the objectives in (1a–c)/(47a–c).

For example, in terms of the impact of non-linguistic factors on the structure of grammar, Paper I demonstrated that the 91.89% rate at which the ending Ice. plural *-ur* is dispersed among feminine nouns motivates the reanalysis of masculine forms in the ending as grammatically feminine. Reanalysis is, therefore, a function of statistical learning, itself a domain-general cognitive process applied to both linguistic and non-linguistic phenomena. However, it is the dispersion of formal attributes such as those associated with the Icelandic masculine class of *-nd-* stems in plural *-ur*, in conjunction with the knowledge that all of these can have biologically male referents, that yields different rates of reanalysis among masculine forms in the same ending.

Thus, the forms of *nd*-stems and other masculines in plural *-ur* align with the medially schematic schema [X*ur*]$_{\text{nom./acc.pl.}}$, which has high cue and category validity for treatment as feminine. However, association of *nd*-stem referents with male biological sex most often impedes such treatment. By the same token, then, alignment with the minimally schematic schema [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$, in addition to medially schematic [X*ur*]$_{\text{nom./acc.pl.}}$ — both of which have high cue validity for assignment of feminine grammatical gender — facilitates category validity for treatment as a member of the feminine X*ó*/*æ*T-microclass specifically. Likewise, the interaction of non-linguistic factors such as schematicity with linguistic factors like semantics are at play in the rate of reanalysis: masculine plural *fætur* is neutral with regard to biological sex, aligns at both levels of abstraction and, therefore, is reanalysed as feminine at a far higher rate than any other masculine form in plural *-ur*.

Continuing with the impact of domain-general cognition on the rate of reanalysis, Paper I demonstrates that the different rates reported are a function of prototype structure — a property of categorisation that stems from statistical learning and schematisation. Indeed, as argued in 5.3.1, perception of structure in the world proves a function of the skewed frequencies with which sets of different attributes are integrated into phenomena considered similar or the same. Thus, while all masculine forms in plural *-ur* are discernible as feminine on alignment with the schema [X*ur*]$_{\text{nom./acc.pl.}}$ — all other things being equal — the schema lacks cue validity for specific class assignment. In this regard, [X*ur*]$_{\text{nom./acc.pl.}}$ serves as a basic-level category, whose level of abstraction is equivalent to that of the non-linguistic category FURNITURE, cf. e.g. basic CHAIR, as discussed in 3.2.4.

However, as the structure of the subtype for feminines in plural *-ur* converges around the X*ó*/*æ*T-microclass, the effect of prototype structure as a mechanism for categorisation

reveals itself as a correlation of the degree of one-to-one correspondence between the schema [X*æ*T*ur*]$_{\text{nom./acc.pl.}}$ and plural ***fætur*** and *ve**tur***, on the one hand, and the rate of reanalysis, on the other. In other words, knowledge of the skewed frequencies with which prototypical features are dispersed among the members of a category is predicated on prior experience with language as stored in memory. Subsequently, this knowledge is extended to new contexts that are perceived as somehow similar to existing knowledge. This conclusion is considered a clear indication of the impact of non-linguistic factors such as frequency of use, dispersion, and schematicity, as well as processes such as statistical learning, analogy, and categorisation on the structure of grammar as a function of domain-general cognition.

In a similar vein, Paper II demonstrates that prior experience with language is stored in memory, even when that experience pertains to low-frequency patterns associated with minimal schematicity. Further, the article indicates that such patterns can display — albeit limited — productivity, a property of language use that is at odds with the "either/or" approach that characterises traditional generative models and those that assume dual-processing. Indeed, the fact that the usage-based cognitive approach assumes graded productivity as a property of language use demonstrates that the theoretical tenets of the approach are in line with the nature of observable language change. Therefore, it is submitted that the approach adopted in the current thesis is more suitable than generative and dual-processing models at accurately accounting for change as a function of usage.

Thus, Papers I and II show that productivity is a graded property of language dependent on mitigating linguistic and non-linguistic factors such as semantics, frequency, and schematicity. For example, the attractive force of the X*ó*/*æ*T-microclass schemas are mediated by semantics and/or pragmatics. Consider the cline according to which forms of Ice. *fótur*, *kók*, and *blók* are treated as members of the microclass in question. While minimal schematicity understandably facilitates assignment of feminine grammatical gender to forms from all three paradigms, this non-linguistic property conspires with semantics in treatment of the borrowing *blók* as feminine only. In other words, due to the nature of the real-world referents of Ice. *blók*, i.e. the noun refers to lowly people of any gender, there is clear grounds for association of the schema [X*ó*T]$_{\text{sg.}}$ with a given real-world function, i.e. form and function converge when the grammatical gender of a word correlates with the biological sex of its referent (PII:215–216). This association, it is argued, facilitates the assignment of feminine grammatical gender (linguistic) on the basis of reference to biological sex (non-linguistic).

Conversely, the paradigm for borrowed *kók* manifests a significant degree of grammatical gender fluidity, by which semantics dictates association between specific forms

of the noun, on the one hand, and neuter and/or feminine grammatical gender, on the other. In light of this property of use, then, association between forms of the noun and their grammatical gender also accords with a cline, by which similarity to a given pragmatic context occasionally motivates use of fem. dat.sg. *kók* instead of neut. dat.sg. *kóki* (see PII:208–209). Furthest along the cline is masculine plural *fætur*, which is the only form of Ice. *fótur* to manifest use as feminine. It is noted in Section 5.2 that this property of use stems from the neutrality of its referent with regard to biological sex, in addition to the overtly masculine marking of most singular forms of the word. It is, therefore, clear that rich memory and the wealth of experience stored therein facilitates categorisation via schema. This domain-general cognitive process is considered the mechanism of language change, itself a function of analogical reasoning, which, in line with Blevins and Blevins (2009), is viewed as the core of all human cognition.

The view of productivity as a graded property of language garners further support from the findings of Papers II and III, both of which account for levelling as a mechanism of analogy. Indeed, Paper II argues for graded membership of the X*ó*/*æ*T-microclass as a function of limited productivity, which, in turn, is a property of low frequency and/or minimally schematic inflection classes. Thus, despite the common treatment of Ice. *kók* as neuter in the *mass* noun sense, the frequency with which the noun occurs with feminine modifiers in the *count* noun sense facilitates levelling of the alternation nom./acc.sg. *kók* ~ dat.sg. *kóki* to syncretic fem. nom./acc./dat.sg. *kók*, where nom./acc.sg. *kók* is considered basic (PII:209). Syncretism in nominative/accusative/dative singular is an exceptionless attribute of strong feminine nouns in Icelandic, knowledge of which is extracted via the process of statistical learning. The deduction of syncretic fem. nom./acc./dat.sg. *kók* suggests that gender assignment is a function of categorisation due to similarity of sg. *kók* to X*ó*/*æ*T-microclass members.

Likewise, Paper III demonstrates that levelling is not "either/or" in nature but, rather, a graded process. Further, levelling is discernibly determined by those conditions in the world that form part of the language user's experience. Indeed, use of individual forms of Far. *vøllur* and *fjørður* was shown not to be constrained by common semantics, cf. that both nouns have topographical referents and can occur as complex place names. Rather, Paper III argues that functionally different forms of the respective paradigms are chosen as basic and that the choice is determined by frequency, which, in turn, reflects language users' perception of real-world conditions. Thus, through the centuries, Far. *vøllur* occurred less frequently in the dative than it does in modern times because its pre-20th century referents tended to be small

and less suitable for performing established functions of the kind that some of its modern referents do.

In other words, Far. *vøllur* was until recently used relatively infrequently in reference to the location of events on account of topographical features of the Faroe Islands. This, it is argued, is the reason that acc.sg. *vøll*, rather than dat.sg. *velli*, was viewed as basic before levelling commenced (PIII:76). Further, as the dispersion of the stem variant *vøll-* was higher among the singular portion of the paradigm, the combined token frequencies of nom.sg. ***vøll****ur* and acc.sg. ***vøll*** facilitated the entrenchment of a form-function pairing through which the form *vøll-* expressed the function *singular*. Such a development is in line with the observable cross-linguistic tendency towards a relation between meaning and form (see Chapter 4).

Likewise, the interaction of token frequency and dispersion is evident through levelling in Far. *fjørður*. There, dat.sg. *firði* is by far the most frequent form of the paradigm in Modern Faroese. Further, given that the noun has not innovated in terms of referents, Paper III argues that its modern-day frequency also proportionately reflects earlier stages of the language. In other words, dat.sg. *firði* is likely also the form that was most entrenched when levelling commenced and, therefore, basic within the paradigm. Should this be correct, we would expect dat.pl. *firðum* to be the earliest yield of levelling. That is, the most entrenched form of the paradigm is more likely to spread to other cells than forms with lower lexical strength. Additionally, we should expect innovative *firðum* to usurp the function of older dat.pl. *fjørðum*, at least in terms of token frequency, as the former contains the most frequent stem variant of the paradigm, which is closely associated with the schematic slot of the schema [*í/úr* FJØRÐUR$_{dat.}$]. Indeed, younger dat.pl. *firðum* is now the most frequent member of the relevant cell (PIII:69).

Further, Paper III demonstrates that deduction and subsequent entrenchment of younger dat.pl. *firðum* established a relation between the stem variant *firð-* and the meaning *plural*, cf. nom./acc.pl. ***firð****ir* (PIII:80). Paper III subsequently argues that, in reaction, language users established a relation between the stem variant *fjørð-* and the meaning *singular*, due to the 2:1 dispersion rate of that variant within the singular portion of the paradigm, cf. nom.sg. ***fjørð****ur* and acc.sg. ***fjørð***. This dynamic, it is argued, yielded innovative dat.sg. *fjørði* as a function of the relation between meaning and form. However, frequency and its impact on cognition once again determines the outcome of reactionary change: Older dat.sg. *firði* is so entrenched as a function of its high token frequency that younger dat.sg. *fjørði* never acquires

the lexical strength needed to take over the function of the former. Indeed, this is supported by the low token frequency of the latter (PIII:69).

In light of the objectives set out above, (1a)/(47a) has been satisfied. Thus, frequency and schematicity — both non-linguistic properties of language use — have been shown to impact domain-general cognitive processes such as statistical learning, entrenchment, and schematisation as applied in linguistic innovation through e.g. categorisation and schematisation. The latter processes feed the former forward as functions of analogy. In support of this assertion, it is possible to refer to the example provided by the cue and category validity of Ice. plural *-ur* for assignment of feminine gender and treatment according to that assignment. Further, the disparate rates at which masculine forms in plural *-ur* are reanalysed as feminine speaks to categorisation by schema, by which entrenchment facilitates the further and gradual elaboration of schemas (PI:11).

As noted in 5.3.1, the perception of structure in the world is a function of the skewed frequencies with which different attributes occur among sets of items perceived as similar. Moreover, this function of skewed frequencies gives rise to prototype structure, itself a domain-independent property of categories. In this connection, Papers I–II demonstrate that linguistic categories are expanded via alignment of linguistic items with schemas at one or multiple levels of abstraction within the same taxonomy. In this connection, it has also been demonstrated that as the level of abstraction incrementally decreases within the relevant taxonomy, the degree of convergence between category members and a category prototype increases in graded fashion. This property of categorisation determines the rate at which Icelandic masculine forms in plural *-ur* are reanalysed as feminine. As this method of expanding categories has been shown to apply in non-linguistic categorisation also, it is concluded that the factors which determine the structure of grammar are domain-general in nature, cf. objective (1b)/(47b).

In order to show that the objective in (1c)/(47c) has been satisfied, it is necessary to refer again to objectives (1a–b)/(47a–b). First, both linguistic and non-linguistic factors have been shown to impact the structure of the synchronic grammar. Secondly, these factors guide the domain-general cognitive processes that are applied in language use. Further, Papers I–III demonstrate that non-linguistic factors such as frequency guide the relevant domain-general processes in real time, during on-line language processing, as each instance of use elaborates and strengthens relevant memory representations for linguistic experience. Through this process, schemas are elaborated to include all new instances of use, thereby increasing the inclusive scope of the existing schema, i.e. its schematicity, and, simultaneously, its potential attractive force as a function of alignment.

In light of the emergent nature of schemas and, by implication, the linguistic categories over which they abstract, analogy applied in real time is considered the mechanism of language use and, therefore, also of change subsequent to entrenchment of the relevant innovation in memory. Thus, without statistical learning, schemas would have no cue validity for class assignment, nor category validity for treatment according to that assignment. Were this the case, Icelandic masculine forms in plural *-ur* would not occasionally be treated as feminine due to the dispersion of the ending. Indeed, the same applies to the borrowings *blók* and *kók* in terms of alignment with the schemas for the X*ó*/*æ*T-microclass. Further, without statistical learning, we should not expect to witness frequency effects in the direction of levelling, based on the choice of most frequent member of the paradigm as basic. However, such effects are observed time and again, cross-linguistically.

It is also possible to posit entrenchment as a function of language use and vice versa. Indeed, positing this give-and-take relation between the two processes is fully in line with the tenet of the usage-based cognitive approach, i.e. that use determines the structure of grammar, while the structure of the grammar informs further use. Thus, it has been demonstrated that statistical learning feeds synchronically into the domain-general cognitive process of entrenchment, yielding ever stronger representation of the relevant experience in memory. On the basis of entrenched knowledge, assignment of items to schematically represented categories occurs. Should categorisation then facilitate reanalyses and/or inflection class assignment (PI–II), or levelling due to entrenchment of a stem variant within a particular context of use (PIII), then language change can be said to have occurred as a function of the domain-general cognitive processes that facilitate language use. Therefore, the current thesis considers objective (1c)/(48c) to have been met also.

In conclusion, then, Papers I–III demonstrate individually that language change is predicated on the process of language use, experience of which is stored in memory and is accessible relative to the degree to which the relevant knowledge is entrenched. In light of this property of language, it stands to reason that the processes which drive the cycle of use and change are domain-general cognitive in nature, rather than attributable to innate cognitive machinery specific to language. It follows, then, that the processes in question are impacted by non-linguistic factors that are nonetheless witnessed through language use and change, such as frequency and its effect on the strength of memory representations for prior experience. Indeed, as argued throughout the current thesis, this property of human cognition is also evident through the application of domain-independent experience as utilised when discerning structure in the world as a function of analogical reasoning.

Paper I

Accounting for different rates of gender reanalysis among Icelandic masculine forms in plural -ur

doi:10.1017/S0332586522000166

Paper II

Tvær blækur labba inn á bar: On limited productivity as graded membership of an Icelandic microclass

doi:10.1075/nowele.00068.mar

Paper III

Um áhrif tíðni á stefnu útjöfnunar: Rannsókn á beygingarþróun færeysku nafnorðanna vøllur og fjørður.

Íslenskt mál og almenn málfræði, 44, 53–86.

Bibliography[74]

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

Albright, Adam. (2008). Explaining universal tendencies and language particulars in analogical change. In Good, Jeff (ed.), *Linguistic Universals and Language Change*, pp. 144–181. Oxford: Oxford University Press.

Albright, Adam. (2009). Modelling analogy as probabilistic grammar. In Blevins, James P., and Juliette Blevins (eds.), *Analogy in Grammar: Form and Acquisition*, 185–213. Oxford: Oxford University Press.

Albright, Adam, and Bruce Hayes. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90, 119–61.

Arge, Jógvan. (1994). *Havnarmenn í Gundadali: Havnar Bóltafelag 1904–1954.* Havnar Bóltafelag.

Audring, Jenny. (2019). Mother or sister? The encoding of morphological knowledge. *Word Structure*, 12, 274–296.

Axelsdóttir, Katrín. (2014). *Sögur af orðum. Sex athuganir á beygingarþróun í íslensku*. Reykjavík: Háskólaútgáfan.

Axelsdóttir, Katrín. (2015). Beyging og merking orðsins *hjalt*. *Orð og tunga*, 17, 95–114.

Árnason, Kristján. (2005). *Íslensk tunga I: Hljóð*. Reykjavík: Almenna bókafélagið.

Árnason, Kristján. (2011). *The Phonology of Icelandic and Faroese.* Oxford: Oxford University Press.

Barðdal, Jóhanna. (2006).[75] Predicting the productivity of argument structure constructions. *Berkeley Linguistics Society*, 32, 467–478.

Barðdal, Jóhanna. (2008). *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam/Philadelphia: John Benjamins.

Barðdal, Jóhanna. (2009). The development of case in Germanic. In Barðdal, Jóhanna, and Shobhana L. Chelliah (eds.), *The Role of Semantic, Pragmatic, and Discourse Factors in the Development of Case.* Amsterdam/Philadelphia: John Benjamins.

Barðdal, Jóhanna. (2011). The rise of dative substitution in the history of Icelandic: A diachronic construction grammar approach. *Lingua*, 121(1), 60–79.

Barlow, Michael, and Suzanne Kemmer. (2000). *Usage-based Models of Language*. Stanford, CA: CSLI.

[74] The bibliography contains all works cited in Papers I–III, in addition to references included in the extended introduction, i.e. Chapters 1–7 of the current thesis.

[75] Included in Paper II as 'Barðdal, Jóhanna. **2012**.' at the request of a reviewer.

Barnes, Michael P. (2001). Faroese Language Studies. *Studia Nordica*, 5. Oslo: Novus Forlag.

Barnes, Michael. P. (2004). Norn — the one-time Scandinavian language of Orkney and Shetland. *Íslenskt mál og almenn málfræði*, 26, 49–82.

Beckner, Clay, and Andrew Wedel. (2010). The roles of acquisition and usage in morphological change. In Kwon, Iksoo, Hannah Pritchett, and Justin Spence (eds.), *Proceedings of the Thirty-fifth Annual Meeting of the Berkeley Linguistics Society February 14-16, 2009*, 1–12. Berkeley, CA: Berkeley Linguistics Society.

Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann. (2009). Language is a complex adaptive system: Position paper. *Language Learning*, 59, 1–26.

Berg, Ivar. (2019). Gender and declension mismatches in West Nordic. In Michela Cennamo, and Claudia Fabrizio (eds.), *Historical Linguistics 2015: Selected Papers from the 22nd International Conference on Historical Linguistics, Naples, 27–31 July 2015*, 97–114. Amsterdam/Philadelphia: John Benjamins.

Bernharðsson, Haraldur. (2004). Um Moldhaugnaháls út í Fjósa og Fjörður. *Íslenskt mál og almenn málfræði*, 26, 11–48.

Bernharðsson, Haraldur. (2005). Ég er, ég vill og ég fær: Þáttur úr beygingarsögu eintölu framsöguháttar nútíðar. *Íslenskt mál og almenn málfræði*, 27, 63–101.

Bernharðsson, Haraldur. (2006). *Gás, gæs* og *Gásir, Gásar*: Brot úr hljóðsögu og beygingarsögu. *Orð og tunga*, 8, 59–91.

Bjorvand, Harald. (1972). Zu den altwestnordischen Pluralendungen *-ar*, *-ir* und *-r* bei femininen Substantiva. *Norwegian Journal of Linguistics*, 26, 195–215.

Bjorvand, Harald. (1975). Altwestnordisch *barar*/*bǫrur*, fpl.: Eine Analyse der analogen Verbreitung der Pluralendung *-ur* der *ōn*-Stämme in den nordischen Sprachen. *Norwegian Journal of Linguistics*, 29, 101–112.

Björnsdóttir, Sigríður Mjöll. (2021). Discovering Gender and Inflection. A view from Icelandic. Ph.D. dissertation. UIT The Arctic University of Norway.

Björnsdóttir, Sigríður Mjöll. (2023). Predicting ineffability: Grammatical gender and noun pluralization in Icelandic. *Glossa*: *a journal of general linguistics*, 8(1), 1–40.

Blevins, James P. (2016). *Word and Paradigm Morphology*. Oxford: Oxford University Press.

Blevins, James P., and Juliette Blevins. (2009). Analogy: An introduction. In Blevins, James P., and Juliette Blevins (eds.), *Analogy in grammar: Form and acquisition*, 1–12. Oxford: Oxford University Press.

Bod, Rens, Jennifer Hay, and Stefanie Jannedy. (2003). Introduction. In Bod, Rens, Jennifer Hay, and Stefanie Jannedy, *Probabilistic Linguistics*, 1–11. Cambridge, MA/London: MIT Press.

Booij, Geert. (2010). *Construction Morphology*. Oxford: Oxford University Press.

Booij, Geert. (2019). The role of schemas in construction morphology. *Word Structure*, 12(3), 385–395.

Booij, Geert, and Jenny Audring. (2018). Partial motivation, multiple motivation: The role of output schemas in morphology. In Booij, Geert (ed.), *The Construction of Words: Advances in Construction Morphology*, 59–80. Cham: Springer.

Brown, Dunstan, and Andrew Hippisley. (2012). *Network Morphology: A Defaults-based Theory of Word Structure*. Cambridge: Cambridge University Press.

Bulloch, Megan J., and John E. Opfer. (2009). What makes relational reasoning smart? Revisiting the perceptual-to-relational shift in the development of generalization. *Developmental Science*, 12(1), 114–22.

Bybee, Joan. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.

Bybee, Joan. (1995). Regular Morphology and the Lexicon. *Language and Cognitive Processes*, 10, 425–455. [Reprinted in Bybee 2007.]

Bybee, Joan. (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, Joan. (2003). Cognitive processes in grammaticalization. In Tomasello, Michael (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, Volume II*, 151–174. New York, NY: Psychology Press.

Bybee, Joan. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733.

Bybee, Joan. (2007). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

Bybee, Joan. (2008). Formal universals as emergent phenomena: The origins of structure preservation. In Good, Jeff (ed.), *Linguistic Universals and Language Change*, 108–121. Oxford/New York, NY: Oxford University Press.

Bybee, Joan. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan. (2015). *Language Change*. Cambridge: Cambridge University Press.

Bybee, Joan, and Clay Beckner. (2015). Language use, cognitive processes and linguistic change. In Bowern, Claire, and Bethwyn Evans (eds.), *The Routledge Handbook of Historical Linguistics*, 503–518. Oxon/New York, NY: Routledge.

Bybee, Joan, and James L. McClelland. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, 22, 381–410.

Bybee, Joan, and Carol Lynn Moder (1983). Morphological classes as natural categories. *Language*, 59, 251–270. [Reprinted in Bybee 2007.]

Bybee, Joan, and Sandra Thompson. (1997). Three frequency effects in syntax. *Berkeley Linguistics Society* 23, 65–85.

Carstairs-McCarthy, Andrew. (2017).[76] Paradigmatic structure: Inflectional paradigms and morphological classes. In Spencer, Andrew, and Arnold M. Zwicky (eds.), *The Handbook of Morphology*, 322–334. Hoboken, NJ: John Wiley and Sons.

Chomsky, Noam. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. (1988). *Language and the Problems of Knowledge*. Cambridge, MA: MIT Press.

Chomsky, Noam. (2017). The Galilean challenge: Architecture and evolution of language. *Journal of Physics: Conference Series*, 880(1). <https://doi.org/10.1088/1742-6596/880/1/012015>

Chomsky, Noam, and Morris Halle. (1968). *The Sound Pattern of English*. New York, NY: Harper and Row.

Christie, Stella, and Dedre Gentner. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development*, 11(3), 356–373.

Cienki, Alan. (2015). Spoken language and usage events. *Language and Cognition*, 7, 499–514.

Clahsen, Harald. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22, 991–1060.

Clausner, Timothy C., and William Croft. (1997). Productivity and schematicity in metaphors. *Cognitive Science*, 21(3), 247–282.

Corbett, Greville. (1991). *Gender*. Cambridge: Cambridge University Press.

Cordes, Anne-Kristin. (2017). The roles of analogy, categorization, and generalisation in entrenchment. In Schmid, Hans-Jörg (ed.), *Entrenchment and the Psychology of Language Learning: How we Reorganise and Adapt Linguistic Knowledge*, 269–288. Washington, DC: American Psychological Association.

Croft, William. (2000). *Explaining Language Change: An Evolutionary Approach*. London: Longman.

Croft, William. (2002). The Darwinization of linguistics. *Selection*, 3(1), 75–91.

Croft, William, and D. Alan Cruse. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cser, András. (2023). Sound patterns, frequency and predictability in inflection. *Acta Linguistica Academica*, 70(3), 397–409.

Davidson, Herbert. 1990. *Han hon den: Genusutvecklingen i svenskan under nysvensk tid* (Lundastudier i nordisk språkvetenskap A45). Lund: Lund University Press.

[76] Mistakenly included in Paper III as 'Carstairs-McCarthy, Andrew. **2015**'.

Dąbrowska, Ewa. (2008). The later development of an early-emerging system: the curious case of the Polish genitive. *Linguistics*, 46(3), 629-650.

Dąbrowska, Ewa. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in Psychology*, 6, 1–17.

De Smet, Hendrik. (2017). Entrenchment effects in language change. In Schmid, Hans-Jörg (ed.), *Entrenchment and the Psychology of Language Learning: How we Reorganise and Adapt Linguistic Knowledge*, 75–99. Washington, DC: American Psychological Association.

Dobson, Annette J. (1973). Estimating time separation for languages. In Dyen, Isidore (ed.), *Lexicostatistics in Genetic Linguistics*, 56-63. Boston: Mouton.

Dressler, Wolfgang U. (1997). On productivity and potentiality in inflectional morphology. *Cross-Language Aphasia Study Network (CLASNET) Working papers*, 7, 3–22.

Dressler, Wolfgang U. (2003). Degrees of grammatical productivity in inflectional morphology. *Italian Journal of Linguistics*, 15, 31–62.

Embick, David. 2015. *The Morpheme: A Theoretical Introduction*. Berlin: Mouton.

Enger, Hans-Olav. (2004). On the relation between gender and declension. *Studies in Language*, 28(1), 51–82.

Enger, Hans-Olav. (2022). Type frequency is not the only factor that determines productivity, so the Tolerance Principle is not enough. *Beiträge zur Geschichte der deutschen Sprache und Literatur*, 144(2),161–187.

Evans, Nicholas, and Stephen C. Levinson. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32(5), 429-448.

Eyþórsson, Þórhallur. (2015). The Insular Nordic experimental kitchen: Changes in case Marking in Icelandic and Faroese. In Whelpton, Matthew, Guðrún Björk Guðsteinsdóttir, Birna Arnbjörnsdóttir, and Martin Regal (eds.), *An Intimacy of Words/Innileiki orðanna. Essays in Honour of Pétur Knútsson/Festschrift til heiðurs Pétri Knútssyni*, 328–352. Reykjavík: Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum/Háskólaútgáfan.

Eyþórsson, Þórhallur, Janne Bondi Johannessen, Signe Laake, and Tor A. Åfarli. (2012). Dative case in Icelandic, Faroese and Norwegian: Preservation and non-preservation. *Nordic Journal of Linguistics*, 35(3), 219–249.

Fertig, David. (2013). *Analogy and Morphological Change*. Edinburgh: Edinburgh University Press.

Finkel, Raphael, and Gregory Stump. (2007). Principal parts and morphological typology. *Morphology*, 17, 39–75.

Frost, Ram, Blair C. Armstrong, Noam Siegelman, and Morten H. Christiansen. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19, 117–125.

Føroyakort. Umhvørvisstovan. <foroyakort.fo>. [accessed 9.7.2021.]

Føroysk orðabók. (1998). Føroysk Fróðskaparfelag og Fróðskaparsetur Føroya.

Gardani, Francesco. (2013). *Dynamics of Morphological Productivity: The Evolution of Noun Classes from Latin to Italian*. Leiden: Brill.

Gardeła, Leszek. (2012). What the Vikings did for fun? Sports and pastimes in medieval northern Europe. *World Archaeology*, 44, 234–247.

Garrett, Andrew. (2008). Paradigmatic uniformity and markedness. In J. Good (ed.), *Linguistic Universals and Langauge Change*, 125–143. Oxford: Oxford University Press.

Gentner, Dedre. (2005). The development of relational category knowledge. In Gershkoff-Stowe, Lisa, and David H. Rakison (eds.), *Building Object Categories in Developmental Time*, 245–275. Mahwah, NJ/London: Lawrence Erlbaum Associates.

Gentner, Dedre, and Jennifer Asmuth. (2019). Metaphoric extension, relational categories, and abstraction. *Language, Cognition and Neuroscience*, 34(10), 1298–1307.

Gentner, Dedre, and Christian Hoyos (2017). Analogy and abstraction. *Topics in Cognitive Science*, 9, 672–93.

Gentner, Dedre, and Arthur B. Markman. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52, 45–56.

Gerken, LouAnn. (2010). Infants use rational decision criteria for choosing among models of their input. *Cognition*, 115(2), 362–366.

Goldberg, Adele E. (2006). *Constructions at Work: The Nature of Generalization in Language*. New York, NY: Oxford University Press.

Greenberg, Joseph H. (1966). *Language Universals: With Special Reference to Feature Hierarchies*. Berlin/New York, NY: Mouton.

Gries, Stephan Th., and Nick C. Ellis. (2015). Statistical measures for usage-based linguistics. *Language Learning*, 65, 228–255.

Guðmundsdóttir, Margrét. (2008). Málbreytingar í ljósi málkunnáttufræði. *Íslenskt mál og almenn málfræði*, 30, 7–52.

Guðmundsson, Valtýr. (1922). *Islandsk Grammatik. Islandsk Nutidssprog*. Copenhagen: H. Hagerup Forlag.

Gunnlaugsson, Guðvarður Már. (2000). Færeyskar málheimildir. In Snædal, Magnús and Turið Sigurðardóttir (eds.), *Frændafundur*, 3, 91–105. Reykjavík: Háskólaútgáfan.

Halford, Graeme S., and Glenda Andrews. (2007). Domain general processes in higher cognition: Analogical reasoning, schema induction and capacity limitations. In Roberts, Maxwell J. (ed.), *Integrating the Mind: Domain General versus Domain Specific Processes in Higher Cognition*, 213–232. New York, NY: Psychology Press.

Hammershaimb, Vencenslaus Ulricus. (1891). *Færøsk Anthologi* I–II. Copenhagen: S.L. Møllers Bogtrykkeri.

Hansson, Gunnar Ólafur. (2007). Productive syncretism in Saami inflectional morphology. In Toivonen, Ida, and Diane Nelson (eds.), *Saami Linguistics*, 91–135. Amsterdam: John Benjamins.

Haspelmath, Martin. (2006). Against markedness (and what to replace it with). *Language*, 42, 25–70.

Hauser, Marc, Noam Chomsky, and William Tecumseh Fitch. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579.

Hay, Jennifer. (2002). From speech perception to morphology: Affix ordering revisited. *Language*, 78, 527–555.

Herce, Borja. (2019). Deconstructing (ir)regularity. *Studies in Language*, 43(1), 44–91.

Hoffmann, Sebastian. (2004). Are low-frequency complex prepositions grammaticalized? On the limits of corpus data — and the importance of intuition. In Lindquist, Hans, and Christian Mair (eds.), *Corpus Approaches to Grammaticalization in English*, 171–210. Amsterdam: John Benjamins.

Hopper, Paul J., and Sandra A. Thompson. (1984). The discourse basis for lexical categories in universal grammar. *Language*, 60, 703–52.

Houghton, Donald E. (1968). Humor as a factor in language change. *The English Journal*, 57, 1178–1181–1186.

isTenTen: Corpus of the Icelandic Web. Scetch Engline. <https://www.sketchengine.eu/isTenTen-Icelandic-corpus/> [Accessed 18.05.2021].

Iversen, Ragnvald. (1972). *Norrøn grammatik*. Oslo: H. Aschehoug and Co.

Íslensk orðabók. (2010). Árnason, Mörður (ed.). Unabridged edition. Reykjavík: Forlagið.

Íslensk orðtíðnibók. See Pind et al. (1991).

Íslenskt textasafn. Stofnun Árna Magnússonar í íslenskum fræðum. <http//corpus.arnastofnun.is> [Accessed 30.07.2022].

Jakobson, Roman. (1939). Signe zéro. *Roman Jakobson, Selected Writings*, 211–219. The Hague: Mouton.

Jakobson, Roman. (1990). Some questions of meaning. In Waugh, Linda R., and Monique Monville-Burston (eds.), *On language: Roman Jakobson*, 315–323. Cambridge, MA: Harvard University Press.

Janda, Laura A. (2002). Cognitive hot spots in the Russian case system. Shapiro, Michael (ed.), *Peircean Semiotics: The State of the Art*, 165–88. New York, NY/Oxford: Berghahn Books.

Janda, Laura A. (2007). Inflectional morphology. In Geeraerts, Dirk, and Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 632–649. New York, NY: Oxford University Press.

Jost, Ethan, and Morten H. Christiansen. (2017). Statistical learning as a domain-general mechanism of entrenchment. In Schmid, Hans-Jörg (ed.), *Entrenchment and the Psychology of Language Learning: How we Reorganise and Adapt Linguistic Knowledge*, 227–44. Washington, DC: American Psychological Association.

Jónsdóttir, Margrét. (1988–1989). Um *ir*- og *ar*-fleirtölu einkvæðra kvenkynsorða í íslensku. *Íslenskt mál og almenn málfræði*, 10–11, 57–83.

Jónsdóttir, Margrét. (1993). Um *ar*- og *ir*-fleirtölu karlkynsnafnorða í nútímaíslensku. *Íslenskt mál og almenn málfræði*, 15, 77–98.

Jónsdóttir, Margrét. (2020). *Epík*, *keramík* og *klassík*: Gerð og beyging fleirkvæðra orða sem enda á *-ík*. *Orð og tunga*, 22, 19–37.

Jónsson, Jóhannes Gísli, and Þórhallur Eyþórsson. (2011). Structured exceptions and case selection in Insular Scandinavian. In Wiese, Heike, and Horst Simon (eds.), *Expecting the Unexpected: Exceptions in Grammar*, 213-241. Berlin: Mouton de Gruyter.

Kay, Paul, and Charles J. Fillmore. (1999). Grammatical constructions and linguistic generalizations: The *what's X doing Y?* construction. *Language*, 75(1), 1–33.

King, Robert D. (1969). *Historical Linguistics and Generative Grammar*. Englewood Cliffs, NJ: Prentice-Hall.

Kiparsky, Paul. (1965). Phonological Change. Ph.D. dissertation. Massachusetts Institute of Technology.

Kiparsky, Paul. (1974). Remarks on analogical change. In Anderson, John Mathieson, and Charles Jones (eds.), *Historical Linguistics*, 257–276. Amsterdam: North-Holland.

Kiparsky, Paul. (2000). Analogy as optimization: 'exceptions' to Sievers' Law in Gothic. In Lahiri, Aditi (ed.), *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology*, 15–46. Berlin/New York, NY: Mouton de Gruyter.

Knudsen, Trygve. (1967). *Kasuslære, vol. 1: Innledning, nominative, akkusative*. Oslo: Universitetsforlaget.

Kodner, Jordan. (2019). Estimating child linguistic experience from historical corpora. *Glossa: a Journal of General Linguistics*, 4(1), 1–14.

Kotovsky, Laura, and Dedre Gentner. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, 67(6), 2797–2822.

Krawczyk, Daniel C., Keith J. Holyoak, and John Hummel. (2004). Structural constraints and object similarity in analogical mapping and inference. *Thinking and Reasoning*, 10, 85–104.

Krawczyk, Daniel C., Keith J. Holyoak, and John Hummel. (2005). The one-to-one constraint in analogical mapping and inference. *Cognitive Science*, 29, 797–806.

Kvaran, Guðrún. (2005). *Íslensk tunga II: Orð*. Reykjavík: Almenna bókafélagið.

Kuryłowicz, Jerzy. (1945–49). La nature des procès dits „analogiques". *Acta Linguistica*, 5, 15–37.

Kürschner, Sebastian, and Damaris Nübling. (2011). The interaction of gender and declension in Germanic languages. *Folia Linguistica*, 45(2), 355–388.

Lahiri, Aditi. (2000). Introduction. In Lahiri, Aditi (ed.), *Analogy, Levelling, Markedness*, 1–14. Berlin: Mouton.

Lakoff, George. (1987). *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.

Lakoff, George. (2018). *Ten Lectures in Cognitive Linguistics*. Leiden: Brill.

Langacker, Ronald. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*, vol. 1. Stanford, CA: Stanford University Press.

Langacker, Ronald. (1988). A usage-based model. In Rudzka-Ostyn, Brygida (ed.), *Topics in Cognitive Linguistics*, 127–161. Amsterdam/Philadelphia: John Benjamins.

Langacker, Ronald. (2008). *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.

Lightfoot, David. (2006). *How New Languages Emerge*. Cambridge: Cambridge University Press.

Magnússon, Ásgeir Blöndal. (1989). *Íslensk orðsifjabók*. Reykjavík: Orðabók Háskólans.

Mańczak, Witold. (1958). Tendances générales des changements analogiques. *Lingua*, 7, 298–325, 387–420.

Markman, Arthur B. (1997). Constraints on analogical inference. *Cognitive Science*, 21, 373–418.

Markman, Arthur B., and C. Hunt Stilwell. (2001). Role-governed categories. *Journal of Experimental and Theoretical Intelligence*, 13, 329–358.

Markússon, Jón Símon. (2012). Eðli *u*-hljóðvarpsvíxla í íslenskri málsögu. MA dissertation. University of Iceland.

Markússon, Jón Símon. (2017). Samband veiklunar og hljóðanvæðingar: Vitnisburður *u*-hljóðvarpvíxla í frum- og vesturnorrænni málsögu. In Hansen, Zakaris Svabo, Anfinnur Johansen, Hjalmar P. Petersen, and Lena Reinert (eds.), *Bók Jógvan: Heiðursrit til Jógvan í Lon Jacobsen á 60 ára degnum*, 263–276. Tórshavn: Fróðskapur.

Markússon, Jón Símon. (2021). Undir áhrifum (orða)gengis. In Axelsdóttir, Katrín, Veturliði Óskarsson and Þorsteinn G. Indriðason (eds.), *Möggubrár heklaðar Margréti Jónsdóttur sjötugri, 21. mars 2021*, 99–104. Reykjavík: Rauðhetta.

Markússon, Jón Símon. (2022a). Tvær blækur labba inn á bar: On limited productivity as graded membership of an Icelandic microclass. *NOWELE*, 75(2), 194–222.

Markússon, Jón Símon. (2022b). Um áhrif tíðni á stefnu útjöfnunar: Rannsókn á beygingarþróun færeysku nafnorðanna *vøllur* og *fjørður*. *Íslenskt mál og almenn málfræði*, 44, 53–86.

Markússon, Jón Símon. (2023a). Accounting for different rates of gender reanalysis among Icelandic masculine forms in plural *-ur*. *Nordic Journal of Linguistics*, 46(3), 331–356.

Markússon, Jón Símon. (2023b). Um hugrænar forsendur fyrir útvíkkun beygingarvíxla: Vitnisburður færeyskra kvenkynsnafnorða með *ar*-fleirtölu. *Ritið*, 23(3), 203–236.

Matras, Christian. (1933). *Stednavne paa de færøske Norðuroyar*. Copenhagen: Aarbøger for Nordisk Oldkyndighed og historie.

Murphy, Gregory L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

OED (*Oxford English Dictionary*). Oxford: Oxford University Press.

ONP (*Ordbog over det norrøne prosasprog*). <https://onp.ku.dk/onp/onp.php> [Accessed 18.1.2022].

Ottósson, Kjartan G. (1992). *The Icelandic Middle Voice: The Morphological and Phonological Development*. Lund: Department of Scandinavian languages, Lund University.

Penn, Derek C., Keith J. Holyoak, and Daniel J. Povinelli. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, 31, 109–178.

Petersen, Hjalmar P. (1992). Skerpingin í føroyskum. In Snædal, Magnús, and Turið Sigurðardóttir (eds.), *Frændafundur*, 11–19. Reykjavík/Tórshavn: Háskólaútgáfan.

Petersen, Hjalmar P. (2009). *Gender Assignment in Modern Faroese*. Hamburg: Verlag Dr. Kovač.

Petersen, Hjalmar P. (2020). *Føroysk míllæra 1. Kyn, orðmyndan og bending*. Tórshavn: Nám.

Petersen, Hjalmar P. and Renata Szczepaniak. (2018). The development of non-paradigmatic linking elements in Faroese and the decline of the genitive case. In Ackermann, Tanja, Horst J. Simon, and Christian Zimmer (eds.), *Germanic Genitives*, 115-149. Amsterdam/Philadelphia: John Benjamins.

Petersen, Hjalmar P., and Laurence Voeltzel. (2024). *Faroese Phonetics andPhonology.* Berlin: Mouton de Gruyter. (To appear.)

Pind, Jörgen (ed.), Friðrik Magnússon, and Stefán Briem. (1991). *Íslensk orðtíðnibók*. Reykjavík: Orðabók Háskólans.

Pinker, Stephen. (1999). *Words and Rules: The Ingredients of Language*. New York, NY: Basic Books.

Popper, Karl. (1959). *The Logic of Scientific Discovery*. London/New York, NY: Routledge.

Posner, Michael I., and Stephen W. Keele. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, 77, 353–363.

Pullum, Geoffrey K., and Barbara Scholz. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review*, 19(1–2), 9–50.

Quinion, Michael. Bloke. *World Wide Words*, 16.6.2004 (updated 22.1.2011). <http://www.worldwidewords.org/qa/qa-blo3.htm> [accessed 17.4.2021].

Rácz, Péter, Clay Beckner, Jennifer Hay, and Janet B. Pierrehumbert. (2020). Morphological convergence as on-line lexical analogy. *Language*, 96(4), 1–36.

Ralli, Angela. (2002). The role of morphology in gender determination: Evidence from Modern Greek. *Linguistics*, 40, 519–551.

Rosch, Eleanor. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–233.

Rosch, Eleanor. (1978). Principles of categorization. In Rosch, Eleanor, and Barbara B. Lloyd (eds.), *Cognition and Categorization*, 27–48. Hillsdale, NJ: Erlbaum.

Rosch, Eleanor, and Carolyn B. Mervis. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson, and Penny Boyes Braem. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.

Ross, Brian H., and Valerie S. Makin. (1999). Prototype versus exemplar models in cognition. In Steinberg, Robert J. (ed.), *The Nature of Cognition*, 205–241. Cambridge, MA: MIT Press.

Rögnvaldsson, Eiríkur. (1981). *u*-hljóðvarp og önnur *a~ö* víxl í nútímaíslensku. *Íslenskt mál og almenn málfræði*, 3, 25–58.

Rögnvaldsson, Eiríkur. (2013). Hljóðkerfi og orðhlutakerfi íslensku. [Electronic edition.] Reykjavík. <https://notendur.hi.is/eirikur/hoi.pdf.>

Schmid, Hans-Jörg. (2017). A framework for understanding linguistic entrenchment and its psychological foundations. In Schmid, Hans-Jörg (ed.), *Entrenchment and the Psychology of Language Learning: How we Reorganise and Adapt Linguistic Knowledge*, 227–44. Washington, DC: American Psychological Association.

Sigurðsson, Einar, Finnur Ágúst Ingimundarson, and Matthew Whelpton. (2022). Samræmi við hulin nafnorð. *Orð og tunga*, 24(1), 21–56.

Sims-Williams, Helen. (2016). Analogical levelling and optimisation: The treatment of pointless lexical allomorphy in Greek. *Transactions of the Philological Society*, 114, 315–338.

Sims-Williams, Helen. (2022). Token frequency as a determinant of morphological change. *Journal of Linguistics*, 58, 571–607.

Sims-Williams, Helen, and Hans-Olav Enger. (2021). The loss of inflection as grammar complication: Evidence from Mainland Scandinavian. *Diachronica*, 38(1), 111–150.

Skjekkeland, Martin. 1997. *Dei norske dialektane*. Kristiansand: Høyskoleforlaget.

Spencer, Andrew. Manufacturing consent over Distributed Morphology. (2019). *Word Structure* 12(2), 208–259.

Stolyarova, Natalia. (2016). Greining á samanburði fallkerfa rússnesku og íslensku. BA dissertation. University of Iceland.

Svavarsdóttir, Ásta. (1993). *Beygingarkerfi nafnorða í íslensku*. Reykjavík: Málvísindastofnun Háskóla Íslands.

Svavarsdóttir, Ásta. (2014). Er virkilega hægt að drekka kók í þremur kynjum? *Vísindavefurinn* 11.12.2014. Archived URL: http://visindavefur.is/svar.php?id=68642 (accessed 12.6.2020).

Sveinsson, Sölvi. (2001). *Íslensk málsaga*. Reykjavík: Iðunn.

Taylor, John R. (2003). *Linguistic Categorisation*. New York, NY: Oxford University Press.

Taylor, John R. (2012). *The Mental Corpus. How Language is Represented in the Mind*. New York, NY: Oxford University Press.

Teldutøka tekstasavn Føroyamálsdeildarinnar. Føroyamálsdeild, Fróðskaparsetur Føroya. <http://www.teldni.fo/tekstasavn/index>. [Accessed 9.7.2021.]

Tiersma, Peter. (1982). Local and general markedness. *Language*, 58, 832–849.

Tímarit.is. <www.timarit.is>. [Accessed 17.5.2022.]

Tomasello, Michael. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1/2), 61–82.

Trudgill, Peter. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford/New York, NY: Oxford University Press.

Tuggy, David. (2007). Schematicity. In Geeraerts, Dirk, and Hubert Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, 82–116. New York, NY: Oxford University Press.

Vikør, Lars S. (2001). *The Nordic Languages: Their Status and Interrelations*. Oslo: Novus Press.

Weyhe, Eivind. (1996). Bendingarmunur í føroyskum málførum. *Íslenskt mál og almenn málfræði*, 18, 71–118.

Weyhe, Eivind. (2012). *Eivindaródn. Greinar 1979–2011*. Tórshavn: Fróðskapur.

Whelpton, Matthew, Drew Trotter, Þórhalla Guðmundsdóttir Beck, Curt Anderson, Joan Maling, Karthik Durvasula, and Alan Beretta. (2014). Portions and sorts in Icelandic: An ERP study. *Brain and Language*, 136, 44–57.

Wiese, Heike, and Joan Maling. (2005). *Biers, kaffi* and *schnaps*: Different grammatical options for restaurant talk coercions in three Germanic languages. *Journal of Germanic Linguistics*, 17, 1–38.

Winters, Margaret E. (2020). *Historical Linguistics: A Cognitive Grammar Introduction*. Amsterdam/Philadelphia: John Benjamins.

Wittgenstein, L. (1978). *Philosophical Investigations*. [English translation by Gertrude Elizabeth Margaret Anscombe]. Oxford: Blackwell.

Wurzel, Wolfgang U. (1984). *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie Verlag.

Yang, Charles. (2016). *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. Cambridge, MA: MIT Press.

Þórhallsdóttir, Guðrún. (1997). Ylgr, heiðr, brúðr: Saga *r*-endingar nefnifalls eintölu kvenkynsorða. In Bragason, Úlfar (ed.), *Íslensk málsaga og textafræði*, 41–56. Reykjavík: Stofnun Sigurðar Nordals.

Þórhallsdóttir, Guðrún. (2007). The dative singular of ō-stems in Old Norse. In Nussbaum, Alan J. (ed.), Verba Docenti: Studies in Historical and Indo-European Linguistics Presented to Jay H. Jasanoff by Students, Colleagues, and Friends, 329–41. Ann Arbor,MI/New York, NY: Beech Stave Press.

Þórólfsson, Björn K. (1925). *Um íslenskar orðmyndir á 14. og 15. öld og breytingar þeirra úr fornmálinu*. Reykjavík: Fjelagsprentsmiðjan. [Reprinted 1987. Reykjavík: Málvísindastofnun Háskóla Íslands.]

Þráinsson, Höskuldur. (2011). Um dauðans óvissan tíma. *U*-hljóðvarp lífs og liðið. *Íslenskt mál og almenn málfræði* 33, 85–107.

Þráinsson, Höskuldur. (2017). *U*-umlaut in Icelandic and Faroese: Survival and death. In Bowern, Claire, Laurence Horn and Raffaella Zanuttini (eds.), *On Looking into Words (and Beyond)*, 99–113. Berlin: Language Science Press.

Þráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen, and Zakaris Svabo Hansen.(2012). *Faroese: An Overview and Reference Grammar*. Tórshavn: FøroyaFróðskaparfelag.