

# 1 Einleitung

Ziel dieser Arbeit ist es, ein Program zu entwickeln, das Lautgesetze auf Wörter anwendet, und die Lautgesetze vom Urindogermanischen in die altgriechischen Dialekte zu sammeln und zu sortieren.

Das junggrammatische Postulat der Ausnahmslosigkeit der Lautgesetze, das fordert, dass Lautgesetze wie mechanisch ablaufen, wirkt wie eine Einladung, tatsächlich einen Mechanismus zu entwickeln, der die Anwendung von Lautgesetzen simuliert. Im 19. Jh. war dies natürlich unmöglich, aber trotz der Computerrevolution des 20. Jh. hat sich offenbar noch niemand daran gemacht, diese eigentlich sehr naheliegenden Idee zu verwirklichen. (TODO: doch, Amir Zeldes ansatzweise)

Eine solche Lautwandelsimulation erfordert, dass die Lautgesetze präzise und in korrekter Chronologie formuliert werden, und erlaubt sowohl, selbige Aufstellung ohne großen Aufwand an beliebig vielen Testwörtern zu überprüfen, als auch belegte Wörter auf ihre Lautgesetzlichkeit hin zu überprüfen.

Das hier vorgestellte Programm `lga` ist der Nachfolger eines Prototyps (komplett in `sed` geschrieben), der primitiver war und einige Mängel aufzeigte. Ebenso ist aber `lga` nur als Prototyp für ein elaborierteres Programm zu verstehen. Im Laufe der Arbeit wird sich herausstellen, welche Probleme, Mängel und weiteren Anforderungen es gibt, die man in einer nächsten Version angehen müsste.

Im ersten Teil dieser Arbeit werde ich `lga` in seinen Designideen und in seiner Funktionsweise vorstellen. Im zweiten Teil werde ich mit Hilfe von `lga` eine relative Chronologie der Lautgesetze, die vom Urindogermanischen in die altgriechischen Dialekte führen, aufstellen. Hierbei ist `lga` prinzipiell natürlich nicht notwendig, aber es ist eine große Hilfe dabei Fehler und Ungenauigkeiten zu finden, die sonst nicht aufgefallen wären.

## 2 Zum Lautwandel allgemein

TODO. Junggrammatiker, Katicic, ...

## 3 lga

### 3.1 Vorüberlegungen

Wenn ein Lautgesetz wirkt, wird ein Laut oder eine Lautgruppe durch eine andere oder nichts ersetzt. Da man Laute phonetisch notieren kann, kann man auch ein Lautgesetz durch Ersetzung von Zeichen durch andere Zeichen beschreiben.

Da Lautwandel meist durch die Lautumgebung bedingt ist und sich oft nicht nur ein Laut sondern Lautklassen verändern (bspw. Stimmloswerdung von stimmhaften Lauten), ist es sinnvoll, eine Sprache zu entwickeln, die Lautgesetze kompakt beschreiben und von einem Computerprogramm verstanden werden kann. Eine naheliegende Wahl ist daher eine Form von *regulären Ausdrücken*, die in vielen Programmen implementiert und weithin bekannt sind. Reguläre Ausdrücke sind zwar für diesen Zweck nicht perfekt geeignet, aber für einen Prototyp ausreichend. Wenn sich gezeigt hat, welche Defizite reguläre Ausdrücke für diesen Anwendungsfall haben, wird man eine besser geeignete Sprache entwickeln und implementieren können.

Als Programmiersprache habe ich die Scheme-Implementation CHICKEN<sup>1</sup> und die *irregex*-Engine für reguläre Ausdrücke<sup>2</sup> gewählt.

### 3.2 Reguläre Ausdrücke

Hier eine (informelle) Beschreibung der hier verwendeten Untermenge regulärer Ausdrücke. Die vollständige Dokumentation für die *irregex*-Engine für CHICKEN findet sich unter dem oben genannten Link.

Die grundlegende Funktion von regulären Ausdrücken ist das Finden (engl. *match*) von Zeichenketten (String) anhand eines Musters. Der gematchte Text kann dann durch einen anderen ersetzt werden, der Teile des Matches wieder aufnehmen kann.

Ein regulärer Ausdruck matcht einen String aus einem Alphabet. Dies ist hier die Menge der Unicode *Codepoints*, nicht die der Unicode *Grapheme*, welche nämlich aus mehreren Codepoints bestehen können.

Ein Zeichen des Alphabets, das kein Metazeichen ist (dazu unten), matcht sich selbst. Metazeichen verlieren ihre spezielle Funktion, wenn ihnen ein Backslash \ vorangeht. Der Ausdruck `foo` matcht also den String `foo` (und auch das `foo` in `foobar`).

Der Punkt `.` matcht ein beliebiges Zeichen. Der Ausdruck `...` matcht also alle Strings mit drei Zeichen, `\\.\\.\\.`  matcht drei Punkte.

Eine in eckigen Klammern stehende Menge an Zeichen matcht eins dieser Zeichen, oder alle außer dieser Zeichen, wenn das erste Zeichen ein Zirkumflex `^` ist. Dabei können auch Zeichenbereiche mit Bindestrich angegeben werden. `[abc]` matcht also `a`, `b` oder `c`, `[^abc]` alle Zeichen außer `a`, `b` und `c`. Alternativ hätte man auch `[a-c]` schreiben können.

Der Zirkumflex `^` und das Dollarzeichen `$` matchen den Anfang bzw. das Ende einer Zeile. `^foo` matcht also `foo` am Anfang der Zeile, `bar$` matcht `bar` am Ende der Zeile und `^quux$` die Zeile, die nur `quux` enthält. Da in `lga` jedes Wort in einer eigenen Zeile

---

<sup>1</sup><http://www.call-cc.org/>

<sup>2</sup><http://wiki.call-cc.org/man/4/Unit%20irregex>

steht, kann man die Zeichen für den Anfang bzw. das Ende des Wortes benutzen.

?, \* und + sind Quantoren und matchen, was ihnen vorangeht, null oder einmal, null bis unendlich mal und ein bis unendlich mal. Der Ausdruck `.[a-d]*x` z.B. matcht optional ein beliebiges Zeichen, dann eins der Zeichen a-d null bis unendlich mal und schließlich mindestens ein x.

Mit Klammern ( ) werden analog zur mathematischen Notation Matches enger gruppiert. So würde `[ab]c` a oder b und dann mindestens ein c matchen (also z.B. `accc` oder `bc`). `([ab]c)` dagegen matcht `[ab]c` mindestens einmal (also z.B. `acbc` oder `bc`).

Neben ihrer Gruppierungsfunktion werden die Submatches innerhalb der Klammern für die Ersetzung gespeichert. `([0-9]+)foo([0-9]+)` würde also zwei Zahlen und `foo` dazwischen matchen. Die Zahlen wären dann in den Submatches 1 und 2 gespeichert und können im Ersetzungstext wieder aufgegriffen werden. Dabei gilt, dass verschachtelte Klammerausdrücke vor Klammern derselben Ebene numeriert werden. Also wären die Submatches in z.B. `(a(b(c))(x))` der Reihe nach `abc`, `b`, `c`, `x`.

Mit dem senkrechten Strich `e0|e1` wird `e0` oder `e1` gematcht (der Senkrechtestrich hat die niedrigste Präzedenz). `foo|bar` matcht also `foo` oder mindestens ein `bar`. `(foo|bar)+` dagegen matcht mindestens ein `foo` oder `bar`.

### 3.3 Funktionsweise

Im Kern arbeitet `lga` mit einer Liste von Wörtern und einer Liste von Lautgesetzen (in der Form von Textersetzung durch reguläre Ausdrücke) und wendet der Reihe nach alle Lautgesetze auf die Liste der Wörter an. Das Resultat sind die Wörter nach Anwendung aller Lautgesetze.

Die Liste der Lautgesetze aber hängt davon ab, von welcher in welche Sprache die Wörter transformiert werden sollen und wird aus einem *Baum* von Lautgesetzen und einem zugehörigen Sprachstammbaum generiert.

Die Liste der Wörter wird vorher durch eine weitere Liste von Regeln, die von der Ausgangssprache abhängt, von einer graphematisch/morphonologischen Form in eine interne phonetische gebracht (was natürlich eine eindeutige Graphie voraussetzt) und am Ende durch noch eine Liste von Regeln, die von der Zielsprache abhängig ist, wieder in eine graphematische Darstellung gebracht.

Die Funktion `(run-list from to words)` generiert die drei Listen von Regeln anhand der Sprachen `from` und `to`, macht aus der Liste von Wörtern `words` aus Effizienzgründen einen einzigen String, in dem jedes Element von `words` in einer eigenen Zeile steht, wendet darauf die drei Regellisten an, und gibt das Ergebnis zurück.

Zunächst wird mit der Funktion `(make-path tree start end)` anhand eines Sprachstammbaums `tree` eine Liste aller Sprachen von der Ausgangs- zur Zielsprache ge-

neriert. Der Aufruf (make-path lang-tree 'uridg 'lesb) würde bspw. die Liste (uridg urgr nwgr aiol lesb) erzeugen.

Der Sprachstammbaum (gespeichert in der globalen Variable lang-tree) definiert das Verwandtschaftsverhältnis von Sprachen. Wenn eine Sprache keine Untersprachen hat, ist sie (in Scheme-Terminologie) ein Symbol (also hat z.B. myk keine Abkömmlinge). Hat eine Sprache Untersprachen, ist sie eine Liste, dessen erstes Element ein Symbol für die Sprache selber ist; die weiteren Elemente sind Untersprachen (wiederum Symbole oder Listen). Ein vereinfachter Stammbaum könnte also bspw. folgendermaßen aussehen:

```
(urgr (nwgr (dor nwdor sardor sdor)
          (aiol boiot (thess wthess othess) lesb))
      (sogr (ach myk ark kypr pamph)
            (ion-att ion att euboi)))
```

Es sei angemerkt, dass die so beschriebenen Verhältnisse nicht streng als genetische Verwandtschaft interpretiert oder zumindest verwendet werden sollten. Att. und ion. mögen zwar auf ein Urion. zurückgeführt werden, aber auch nach ihrer Aufspaltung teilen sie noch gewisse Entwicklungen. Ebenso gibt es gemeingriechische Entwicklung die nach-urgriechisch sind. Eine präzisere Trennung von Verwandtschaft und "Sprachgruppen" würde aber das Modell möglicherweise zu sehr verkomplizieren. Die beste Möglichkeit ist es womöglich, auf die Vorsilbe *ur-* zu verzichten und bspw. *gr* gleichermaßen für urgriechisch und gemeingriechisch zu verwenden (ebenso wäre statt *urion* besser *ion-att* zu verwenden).

Der Lautgesetzbaum ist eine Liste von vier möglichen Elementen, nämlich Funktionen, die einen String als Argument nehmen und einen String zurückgeben (also Funktionen, die Lautgesetze anwenden), Symbole, die eine erreichte Sprachstufe markieren, und zwei Arten von Listen des Typs (br|sub lang-lst rule-tree), deren zweites Element eine Liste von Sprachen ist, für die sie gelten (bzw. nicht gelten, wenn das erste Element not ist), und deren drittes Element ein weiterer Lautgesetzbaum ist. *br* (für *branch*) leitet einen neuen Sprachzweig ein, der durch *rule-tree* definiert ist (d.h. dass alle Regeln nach einem erfolgreichen *br* ignoriert werden). *sub* (für *subrules*) funktioniert ähnlich, macht aber nach Abarbeitung von *rule-tree* nach der Regel weiter. Dies impliziert, dass es sich hier bei *rule-tree* nicht um einen echten Baum handelt. *sub* ist nur dazu gedacht eine Liste von Lautgesetzen bedingt anzuwenden. Im Griechischen ist dies besonders nützlich, da zwar alle Dialekte ähnliche Entwicklungen machen, diese sich aber im einzelnen unterscheiden. Mit *sub* kann man einfach Regeln für bestimmte Dialekte oder Dialektgruppen definieren ohne für jeden Dialekt einen eigenen Sprachzweig haben zu müssen.

Die Funktion `(make-rules tree path)` gibt anhand des Lautgesetzbaums `tree` eine Liste von Lautgesetzen zurück, die die Sprachentwicklung in `path` (von `make-path` erzeugt, s.o.) reflektiert. Von `(apply-rules rules words)` wird dann diese Liste `rules` auf die Wörter `words` angewandt.

Die Lautgesetzfunktionen werden von der Funktion `(s pattern . subst)` erzeugt, die aus einem String `pattern` einen regulären Ausdruck macht und eine Funktion zurückgibt, die `irregex-replace/all` auf diesen regulären Ausdruck und die Ersetzungen in der Liste `subst` anwendet. Matcht der reguläre Ausdruck einen Teil der Eingabe, wird diese durch `subst` ersetzt, dessen Elemente Strings, Zahlen als Indizes der Submatches, sowie Funktionen, die anhand eines Matches Strings zurückgeben, sind. `(irregex-replace/all "(foo)(bar)" "foobar" 2 (lambda (m) (string-reverse (irregex-match-substring m 1)))) "quux"` z.B. würde den String "foobar" durch "baroofquux" ersetzen.

Die oben beschriebenen regulären Ausdrücke sind für die Modellierung von Lautgesetzen nicht ausreichend. Neben anderen Unzulänglichkeiten, die sich im Laufe der Arbeit herausstellen werden, gibt es eindeutig Probleme, mit Lautklassen umzugehen, und noch allgemeiner, Laute durch Zeichen zu kodieren.

Wenn jeder Laut durch genau ein Zeichen bzw. einen Unicode Codepoint kodiert wird, gibt es der Lautkodierung keine Probleme. Da es allerdings sinnvoll ist, einige Laute mit mehreren Codepoints zu kodieren, kann es zu Problemen kommen. Will man bspw. ein Lautgesetz `(s "b" "p")` formulieren, so hätte dies die Nebenwirkung, dass auch  $b^h$  zu  $p^h$  würde, was im allgemeinen Fall unerwünscht ist. Wenn man Laute wie  $b^b$  oder  $g^w$  mit einem einzigen Codepoint repräsentieren würde, hätte man dieses Problem natürlich zwar nicht, jedoch bietet Unicode für solche Vorhaben keine definierten Codepoints. Dafür müsste man die Private Use Area benutzen, wodurch die Kodierung jedoch von der Schriftart abhängig wird, eine unschöne Lösung. Eine Möglichkeit, dieses Problem zu umgehen, ist, Polygraphen sowohl in den zu transformierenden Wörtern als auch in den die Lautgesetze beschreibenden regulären Ausdrücken durch beliebige unbenutzte Codepoints zu ersetzen und am Ende die umgekehrte Ersetzung durchzuführen, so dass die nur intern verwendeten Codepoints nirgendwo graphisch auftauchen. In der hier entwickelten Version von `lga` wird das Problem aus Effizienzgründen jedoch noch nicht gelöst. Stattdessen müssen die Lautgesetze vorläufig besonders vorsichtig formuliert werden, der obere Fall also als `(s "b([h])" "p" 1)`.

Bei der Formulierung von Lautgesetzen ist es ferner sinnvoll, Laute in Lautklassen zusammenzufassen um eine konzise und leicht abstrahierende Darstellung zu haben. Der POSIX Standard definiert zwar Zeichenklassen wie `[:digit:]` für Ziffern

oder [:lower:] für Kleinbuchstaben, aber die Bedürfnisse, die man als Linguist hat, sind damit nicht abgedeckt, da sich zur Laufzeit Zeichenklassen weder neu definieren noch verändern lassen. Dieses Problem lässt sich vorläufig lösen, indem die regulären Ausdrücke, bevor sie von `string->irregex` in eine interne Form kompiliert werden, noch verändert werden, so dass die tatsächlich im Quelltext vorkommenden regulären Ausdrücke erst zu echten regulären Ausdrücken verarbeitet werden, die von der `irregex`-Engine verstanden werden. Konkret funktioniert dies so, dass in einer assoziativen Liste Lautklassennamen (als Symbole) mit Strings assoziiert werden und bei der Vorbearbeitung der regulären Ausdrücke Strings der Form `<lautklassenname>` durch den mit dem Symbol `lautklassenname` assoziierten String ersetzt werden. Die Funktion `(set-class key value)` assoziiert den Namen `key` mit dem String `value` und `(l key)` (kurz für *lookup*) findet den zu `key` passenden String in der assoziativen Liste.

So kann man bspw. mit `(set-class 'kurz-vok "a|e|i|o|u")` und `(set-class 'lang-vok "ā|ē|ī|ō|ū")` Kurz- und Langvokale definieren, mit `(set-class 'vok (s+ (l 'kurz-vok) "|" (l 'lang-vok)))` dann die Menge aller Vokale (`s+` ist eine Abkürzung für `string-append` zum Verketteten von Strings) und diese in einem Lautgesetz wie `(s "<vok>" "a")` verwenden.

Die Lautklassen sind jedoch zur Laufzeit nicht vernünftig veränderbar, da sie zur Zeit der Auswertung von `s` in den regulären Ausdruck eingefügt werden, man die Definition aber wohl am ehesten aus dem Lautgesetzbaum heraus verändern wollen würde (Lautklassen sollten also dynamisch statt lexikalisch gebunden sein). Hier wird nur eine neuentwickelte Lautgesetzengine wirklich Abhilfe schaffen können.

Es gibt jedoch noch ein weiteres Problem. So kann man nicht ohne weiteres z.B. einen Langvokal durch seinen entsprechenden Kurzvokal oder einen silbischen durch sein unsilbisches Pendant ersetzen. Beispiele für solche Vorgänge gibt es genug und es ist eine Methode notwendig, mit der man diese Fälle unkompliziert ausdrücken kann. In einer späteren Version wäre vielleicht eine Ersetzung wie `(s "<lang-vok>" "<kurz-vok>")` wünschenswert, in der aktuellen Version können jedoch auch Funktionen, die auf Submatches angewandt werden, denselben Effekt erzielen. Als Ersetzungsargumente kann man an `irregex-replace/all` neben Strings und Zahlen, die das entsprechende Submatch bezeichnen, auch Funktionen übergeben, die auf das aktuelle Match angewandt werden und einen String zurückgeben. Um also z.B. einen Langvokal in ein Kurzvokal zu verwandeln, braucht es nur eine Funktion, die auf ein Submatch eine Reihe von regulären Ausdrücken anwendet und das Ergebnis zurückgibt. Eine Funktion, die reguläre Ausdrücke auf einen String anwendet, gibt es ja schon: `apply-rules`. Diese wird von der Funktion `(match-rulelist rules)` benutzt, die eine

Liste von Ersetzungsfunktionen `rules` auf ein Submatch angewendet. Dazu gibt sie eine Funktion zurück, die einen Submatchindex `i` bindet und eine Funktion zurückgibt, auf die `irregex-replace/all` angewandt werden kann. Die Vokalkürzung könnte somit folgendermaßen definiert werden:

```
(define kuerzung
  (match-rulelist
    (list (s "ā" "a")
          (s "ē" "e")
          (s "ō" "o")
          (s "ī" "i")
          (s "ū" "u")))))
```

Ein Lautgesetz, das alle Langvokale kürzt, könnte dann als `(s "(<lang-vok>)" (kuerzung 1))` formuliert werden.

Die eingangs erwähnten Regellisten zur Umwandlung von einer graphematisch/morphonologischen in eine phonetische und von einer phonetischen in eine graphematisch/phonologische Darstellung sind unter ihrem Sprachnamen ebenfalls in assoziativen Listen abgelegt.

Hiermit ist die Funktionalität von `lga` weitestgehend erläutert. Eine exakte Beschreibung kann natürlich nur der Quelltext selber liefern.

## 4 Modellierung der Lautentwicklung des Altgriechischen

### 4.1 Vorbemerkungen

Vor der Beschreibung der Lautentwicklung noch einige Anmerkungen, die nicht speziell das Griechische betreffen. Aufgrund von Fehlern in der Implementation von `irregex` können in regulären Ausdrücken keine höheren Unicodezeichen in `[]`-Sets benutzt werden. Da diese aber notwendig sind, um gewisse Zeichen auszuschließen – insbesondere <sup>h</sup> zur Unterscheidung von Aspiraten – müssen diese aus dem ASCII-Vorrat stammen. Aus dem Grund wird für die Aspiration `!` statt <sup>h</sup> verwendet. Eigentlich wäre wie schon oben beschrieben für Aspiraten oder sonstige Laute, die mit mehreren Codepoints ausgedrückt werden, ein einziger Codepoint oder eine Möglichkeit, mehrere Codepoints als ein Zeichen zu behandeln, praktischer und angemessen. Für die Labiovelare wurden im Gegensatz zu den Velaren Großbuchstaben verwendet, um dieses Problem zu vermeiden und auch die silbischen Resonanten *ῥ**ῖ**ῗ* werden mit den Großbuchstaben `RLMN` ausgedrückt. Die Halbvokale *ιυ* werden mit `yw` und die Laryngale *h*<sub>[123]</sub> mit `H[123]` bezeichnet. Akzente werden in der graphematischen Schreibweise als combining diacritics geschrieben und intern mit `'` und `~` bezeichnet.

Die Umwandlung dieser graphematischen Schreibungen in die interne (mehr oder weniger phonetische) Darstellung geschieht in der Liste der phonologischen Regeln für uridg (in `rules.scm`).

Als weitere phonologische Regeln für das Uridg. sind Laryngalfärbung (\**e* neben \**h<sub>3</sub>* jedoch noch von \**o* verschieden, was allerdings für das Gr. belanglos ist) und Stimmhaftwerdung von \**s* neben stimmhaften Konsonanten beschrieben. Weitere phonologische Regeln müssen ggf. ergänzt werden. Lautgesetze wie STANGS Gesetz, die vor dem Erreichen des Sprachzustandes, der mit uridg bezeichnet wird, durchlaufen wurden, wurden nicht berücksichtigt und die Rekonstrukte bzw. Transponate, die als Eingabe dienen, sollten einen entsprechenden Lautstand vorweisen (um bei STANGS Gesetz zu beleiben wäre also \**djēm* nicht \**djeum* uridg). Insbesondere enthält uridg den *þ*-Laut, dessen Herkunft und Phonetik hier ebenfalls nicht weiter hinterfragt wird.

Im Folgenden werden die Lautgesetze vom Urindogermanischen ins Griechische in drei Etappen aufgestellt. Zunächst wird die Lautentwicklung bis ins Urgriechische modelliert, dann bis und in mykenischer Zeit und schließlich in die Dialekte der alphabetischen Zeit. Die dazu verwendete Literatur ist im wesentlichen Rix 1992, Bartoněk 1991, Bartoněk 2003, Sihler 1995, Buck 1910.

Eine (relative) Chronologie aller Lautgesetze lässt sich nicht immer mit Sicherheit aufstellen, da viele Lautgesetze kaum mit anderen Lautgesetzen interagieren oder mehrere Entwicklungen denkbar sind, so dass man gerade in der frühesten Zeit, in der das Griechische noch nicht belegt ist, in diesen Fällen wenig mehr als nur raten kann. Hier werden dann die Lautgesetze entweder mit ähnlichen anderen Gesetzen oder nach Gefühl einsortiert.

Auch ist die Aufstellung hier keinesfalls als endgültig anzusehen. Besonders Entwicklungen, die in die spätere dialektale Zeit fallen wie Digammaschwund und Kontraktionen die zum Teil nach Einsetzen der Überlieferung stattfinden, sind zu unübersichtlich und unklar, um sie hier angemessen zu modellieren ohne den Rahmen der Arbeit zu sprengen (Schwierigkeiten werden im Text genannt). Aber auch über sicherere Entwicklungen lässt sich im Detail fast immer streiten. Sowohl das Programm `lga` als auch die Formulierung der Lautgesetze sind somit als vorläufiges Ergebnis oder Gerüst zu sehen, auf dem weitere Arbeiten aufbauen sollen.

## 4.2 Das Lautsystem des Urindogermanischen

Das hier angenommene Lautsystem des Uridg. besitzt folgende Laute. Umstrittene Laute stehen in Klammern. Teilweise ist nur der Phonemstatus umstritten, aber Phoneme sind für Lautgesetze ohnehin uninteressant. Andererseits sind die hier aufgeführten Laute natürlich immer noch eine Abstraktion über die tatsächlichen Laute,



was aber in der Natur der Rekonstruktion liegt und auch aus pragmatischen Gründen kaum anders gehandhabt werden könnte.

Vokale:  $*a, *ā, *e, *ē, *o, *ō, *i, (*ī), *u, (*ū)$

Silbische Resonanten:  $*r, *l, *m, *n$

Resonanten und Halbvokale:  $*r, *l, *m, *n; *i, *u$

Verschlusslaute:

Labiovelare  $*k^w \quad *g^w \quad *g^{wb} \quad (*k^{wb})$

(Rein)velare  $*k \quad *g \quad *g^b \quad (*k^b)$

Palatale  $*\hat{k} \quad *\hat{g} \quad *\hat{g}^b \quad (*\hat{k}^b)$

Dentale  $*t \quad *d \quad *d^b \quad (*t^b)$

Labiale  $*p \quad *b \quad *b^b \quad (*p^b)$

Frikative und sonstige:  $*s, *z, *h_1, *h_2, *h_3, (*\beta)$

$*z$  ist Allophon von  $*s$  vor stimmhaften Konsonanten. Beide werden aber der Einfachheit halber in  $\text{ḷga}$  als  $*s$  repräsentiert.

Die genaue Artikulation der Laryngale  $*h_1, *h_2, *h_3$  ist nicht klar. Meist werden Frikative vermutet, neben denen ggf. Sprossvokale eingefügt werden.

Der Laut  $*\beta$  ist umstritten, sowohl seine Existenz als auch seine Artikulation (vielleicht dentale Affrikata?).

### 4.3 Urindogermanisch bis Urgriechisch

Die Lautveränderung, die ins Urgr. führen, sind grob gesprochen die Kentumvertretung der Tektale, die Stimmloswerdung der Mediae aspiratae, die Entwicklung von  $*\beta$ , jegliche Laryngalentwicklungen, die Entwicklung der silbischen Resonanten (teilweise aber auch erst dialektal), einige Okklusivassimilationen, die Entwicklung von  $*s$  und die Anfänge diverser Palatalisierungen.

**1 Kentum** Das Griechische ist eine Kentumssprache, es fallen also die uridg. Palatale mit den Velaren zusammen (Rix §92–94). Da das Griechische keinerlei Anhaltspunkte dafür bietet, dass es je einen Unterschied zwischen diesen Lauten gab, wird der Zusammenfall früh datiert.

$*\hat{k}, *\hat{g}, *\hat{g}^b, (*\hat{k}^b) > *k, *g, *g^b, (*k^b)$

uridg.  $*\hat{k}$ :  $*\hat{k}m̥tóm$ : gr. ἑκατόν vs. ai.  $\acute{s}atám$  ‘hundert’

uridg.  $*k$ :  $*leukós$ : gr. λευκός vs. ai.  $rocás$  ‘leuchtend’

uridg.  $*\hat{g}$ :  $*h_2égonti$ : att. ἄγουσι vs. av.  $azənti$  ‘sie führen’

uridg.  $*g$ :  $*gerh_2nós$ : gr. γερωνός vs. lit.  $gervė$  ‘Kranich’

uridg.  $*\hat{g}^b$ :  $*ueg̃h-/uog̃h-$ : att. ὄχος ‘Wagen’ vs. av.  $vazāmi$  ‘ich fahre’

uridg.  $*g^b$ :  $*h_3mig^b-/h_3moig^b-$ : gr. ὀμίχλη ‘Nebel’ vs. ai. *meghás* ‘Wolke’

**2 Thorn**  $*p$  (genauer Lautwert unklar) tritt nach Velaren und Labiovelaren auf. Nach Tenuis entspricht  $*t$ , nach Media aspirata  $*d^b$  (Rix §81).

$*p > *t \setminus \{\text{Tenuis}\}_-$

uridg.  $*k^b$ :  $*tek^bō(n)$ : gr. τέκτων vs. ai. *tákṣā* ‘Zimmermann’

$*p > *d^b \setminus \{\text{Media asp.}\}_-$

uridg.  $*g^{wb}$ :  $*ḡg^{wb}pitom$ : gr. ἄφθιτον vs. ai. *ákṣitam* ‘unvergänglich’

**3  $*MA > TA$**  Mediae aspiratae werden zu Tenuis aspiratae (Rix §94). Aufgrund der Nähe zum Makedonischen und Phrygischen, in denen der Reflex vermutlich Media ist, ist dieser Wandel vielleicht später anzusetzen. Spätestens im Myk. ist die Stimmlosigkeit belegt.

$*b^b, *d^b, *g^b, *g^{wb} > *p^b, *t^b, *k^b, *k^{bw}$

uridg.  $*b^b$ :  $*b^béronti$ : dor. φέροντι vs. ai. *bháranti* ‘sie tragen’

uridg.  $*d^b$ : zu  $*d^bēh_1-$ : att. τίθημι vs. ai. *dadhámi* ‘ich setze/stelle/lege’

uridg.  $*g^b$ : zu  $*steig^b-$ : gr. στείχω vs. got. *steigan* ‘steigen’

uridg.  $*g^{wb}$ : zu  $*g^{wben}-$ : gr. φόνος ‘Mord’ vs. ai. *g^bhnánti* ‘sie erschlagen’ (zu  $*k^{wb} > *p^b$  s. TODO)

**4  $*(H)i-$**  Da es im Gr. scheinbar zwei Reflexe von  $*i-$  gibt, versucht man mit der Laryngaltheorie die beiden Reflexe als  $*H_i-$  und  $*i-$  zu unterscheiden. Welcher Anlaut zu urgr.  $*dz-$  verschärft wurde, ist unklar – nach Rix §68,80e ist es  $*H_i-$  – und nach dem Datenmaterial kann für beides argumentiert werden, wenn man aber ὑγίης ‘gesund’ aus  $*h_2iu-g^wih_3-és$  ‘langes Leben habend’ herleiten will ( $*h_1su-$  würde εὖ- ergeben), darf man  $*H_i- > *i-$  vermuten. Auch lautlich plausibel ist eine frühe Verschärfung von uridg.  $*i-$  zu urgr.  $*dz-$  (vielleicht über Zusammenfall mit ererbtem  $*d_i-$ , das ebenfalls zu  $*dz-$  wird), danach Wegfall aller Laryngale (also keine Vokalisierung wie vor  $*u$ ), wodurch  $*H_i-$  als  $*i-$  stehenbleibt, wie es im Frühmyk. belegt ist.

$*i- > *d_i-$

uridg.  $*i-$ :  $*iugóm$ : gr. ζυγόν vs. ai. *yugám* ‘Joch’

$*H_i- > *i-$

uridg.  $*h_2i-$ : s.o. als  $*h_2iu-$  ‘langes Leben ( $\rightarrow$  lange Zeit  $\rightarrow$  Zeit)’ wegen Vollstufe  $*h_3ei-$  in z.B. hom. αἰεί ‘immer’ <  $*h_2eiueśi$  ‘zu allen Zeiten’ zu rekonstruieren.

**5 Laryngale** Die Entwicklung der Laryngale folgt Rix §79-85. Teilweise sind die Lautgesetze weniger explizit als bei Rix formuliert, um eine lautlich plausiblere Entwicklung zu modellieren. So wird z.B. intervokalischer Laryngal nach  $*i$  oder  $*u$  nicht

zum Gleitlaut sondern dieser entsteht erst als Hiattilger nach dem Laryngalschwund. Der Schwund der Laryngale dürfte auch eher früh passiert sein. Der Wandel, der hier als  $*H > *V$  geschrieben ist, ist als  $*h_1, *h_2, *h_3 > *e, *a, *o$  zu verstehen.

$*H > *V \setminus C\_ \#$

$*(i, u)H > *(i, u)V \setminus CC\_ \#$

$*(i, u)H > *(\dot{i}, \dot{u})V \setminus C\_ \#$

$*VH > *\bar{V} \setminus \_ (C, R, \#)$

$*H > *V \setminus (\#, C) \_ C$

$*_{\circ}h_2 > *na \setminus \_ \#$

$*HR_{\circ} > *VR \setminus C\_$

$*\acute{R}_{\circ}H > R\bar{V} \setminus \_ C$

$*R_{\circ}H > \acute{V}RV \setminus \_ C$

$*H(u, R_{\circ}) > *V(\dot{u}, R) \setminus \# \_$

$*H > \emptyset$

**6**  $*_{\circ}ue_{\circ} > ue_{\circ}i$   $*_{\circ}ue_{\circ}$  wird zu  $*ue_{\circ}i$  dissimiliert, wie  $\epsilon\acute{\iota}\pi\omicron\nu < *e-ue-uk^w-om$  (vgl. ai. *avocam*) lehrt. Dies muss ferner vor 7 passieren, wie  $\epsilon\acute{\iota}\pi\omicron\nu$  ebenfalls zeigt.

$*_{\circ}u > *_{\circ}i \setminus ue\_$

**7** Labiovelar > Velar Neben  $*_{\circ}u$ ,  $*_{\circ}i$  und vor  $*_{\circ}i$  fallen die Labiovelare mit den Velaren zusammen (Rix §97).

$*k^w, g^w, k^{wb} > *k, g, k^b \setminus (u, \bar{u}, \dot{u}) \_ ; \_ (u, \bar{u}, \dot{u}, \dot{i})$

**8** Silbische Resonanten Die Entwicklung der silbischen Resonanten folgt Rix §75 und §76. Für das Urgr. wird ein Sprossvokal  $*\partial$  angesetzt, welcher sich dialektal zu *a* oder *o* entwickelt.

Genauere Datierung unklar.

$*N_{\circ} > *aN \setminus (V, \dot{V})$

$*N_{\circ} > *\partial$

$*R_{\circ} > *\partial R \setminus (\#, V, \dot{V})$

$*R_{\circ} > *R\partial$

**9** Nasal vor Okklusiv Vor Okklusiven wird der homorgane Nasal realisiert (Rix §78). Gegen Rix jedoch nicht  $*ms > *ns$  wegen  $\epsilon\nu\epsilon\iota\mu\alpha < *enemsa$ . Hier kann *m* zwar

analog wieder eingeführt worden sein, aber ohne ein Beispiel, das bei Rix fehlt, ist die Regel zunächst unnötig. Die Nasalassimilation passiert auch später noch und dürfte tatsächlich wohl über lange Zeit eine synchrone phonologische Regel sein.

$*N > *m \setminus \{\text{Labial}\}$

$*N > *n \setminus \{\text{Dental, Velar, Labiovelar}\}$

**10**  $*m_i > *n_i$  Rix §77. Datierung unklar. Vielleicht nach  $*m > *am$ , aber  $*m > *n \setminus *i$  ist ebenfalls als zusätzliche Regel denkbar.

$*m > *n \setminus i$

**11** Benachbarte Vokale nach Laryngalschwund (1. Kontraktion) Kontraktion von  $*e, *a, *o$ , zwischen denen Laryngal geschwunden ist (Rix §81). Hier entsteht zuerst der Intonationsunterschied zwischen Akut und Zirkumflex. Trägt der erste Vokal den Akzent, ist das Kontraktionsprodukt zirkumflektiert. Trägt ihn der zweite, ist es akutiert. Datierung unklar.

$*VV > *\bar{V} (*V = *a, *e, *o)$

Nach Vokal werden  $*i, *u$  zu Halbvokalen (es entstehen also Diphthonge). Im umgekehrten Fall entsteht ein Gleitlaut zwischen den beiden Lauten.

$*i, u > *i, u \setminus V\_$

$*i, u > *ii, uu \setminus \_V$

**12** Okklusivassimilation Okklusive werden an folgende Okklusive in Aspiration und Stimmhaftigkeit assimiliert (Rix §106a).

Hier dürfte es sich wohl eigentlich um eine phonologische Regel handeln.

$*\{\text{Okklu.}\} > *\{\text{Media}\} \setminus \{\text{Media}\}$

$*\{\text{Okklu.}\} > *\{\text{Tenuis}\} \setminus \{\text{Tenuis}\}$

$*\{\text{Okklu.}\} > *\{\text{Aspirata}\} \setminus \{\text{Aspirata}\}$

**13**  $*K_u > *(K)K$  Velar +  $*u$  wird zum Labiovelar, zwischen Vokalen als Geminat (Rix §104), um das Silbengewicht zu erhalten (vielleicht nur nach Kurzvokal?).

$*K_u > *K^wK^w \setminus V\_V$

$*K_u > *K^w$

**14** Beseitigung von Geminaten Das Gr. muss eine Zeit lang Geminaten vermieden haben. Entweder generell wie das Uridg. – dann müsste diese Regel vor 13 kommen und dort die Geminatbildung nur nach Kurzvokal passieren – oder nur im Silbenanlaut (also

nur nicht nach Kurzvokal) wie αἰπόλος < \*aḱ<sup>w</sup>olos < \*aḱg-k<sup>w</sup>olos zeigt. Weitere Beispiele könnten die Verhältnisse klarer machen.

Datierung unklar. Vermutlich nach 13 und sicher vor 22.

\*C: > \*C \ . \_

**15** Assimilation über s Eine Tenuis hinter \*s wird an eine vor \*s stehende Media oder Tenuis aspirata in ihrer Artikulationsart angeglichen (Rix §105).

Auf jeden Fall vor 17 wegen πῶσχω < \*p<sup>h</sup>at<sup>h</sup>skō, λέσχη < \*lek<sup>h</sup>skā, μίσγω < \*miskō.

\*{Tenuis} > \*{Media} \ {Media}s \_

\*{Tenuis} > \*{Aspirata} \ {Aspirata}s \_

**16** CsC > sC Steht \*s zwischen zwei gleichen Konsonanten, schwindet der erste (Rix §104). Welche Konsonanten das im einzelnen sind, scheint nicht ganz klar zu sein.

\*C<sub>1</sub> > \*∅ \ \_C<sub>2</sub> (C<sub>1</sub> = C<sub>2</sub>)

**17** Cs Vor s werden Okklusive zu Tenuis (Rix §105).

Auf jeden Fall nach 15 wegen μίσγω.

\*{Okkl.} > \*{Tenuis} \ s

## urgr1

**18** \*s > \*h \*s > \*h vor Vokal oder Resonant oder zwischen Vokalen (Rix §86,89). Vor \*m ist \*s jedoch (sogar im selben Wort) teilweise erhalten.

Datierung unklar.

**19** \*s neben \*j Nach Rix §89g ist \*s neben \*j über \*h assimiliert worden. Die Beispiele belegen jedoch nur eine Assimilation bei \*sḱ und der Weg über \*h scheint nur aus systematischen Gründen gewählt zu sein.

Tatsächlich gibt es nämlich Probleme wenn man die Dat.Pl. Endung -οισι lautgesetzlich aus \*-oḱsi erklären möchte. Daher wird hier auf die Annahme einer Assimilation bei \*j<sup>s</sup> verzichtet.

Datierung unklar.

**20** OSTHOFF OSTHOFFS Gesetz kürzt Langvokale vor tautosyllabischem Resonanten (Rix §64). Im Auslaut ist die antevokalische Sandhivariante verallgemeinert und Langvokale bleiben erhalten. Nach Laryngaldehnung (5) und vor Okklusivverlust im Auslaut (21) zu datieren, wie z.B. hom. ἔσταν < \*estānt < \*e-steh<sub>2</sub>-nt zeigt.

**21 Konsonanten im Auslaut** Wortauslautende Okklusive gehen verloren (Rix §100). Dies muss nach der Vokalisierung der silbischen Liquiden passiert sein, da  $\acute{\upsilon}\pi\acute{o}\delta\rho\alpha < *upodrak < *upodr\acute{k}$  den auslautenden Konsonanten voraussetzt. Außerdem nach OSTHOFF (20) wie dort gezeigt.

Auslautendes  $*m$  erscheint als  $n$  (Rix §77).

Genauere Datierung in beiden Fällen unklar.

**22 Palatalisierung** Die Anfänge der Palatalisierungen liegen wohl in urgr. Zeit und die genaue Entwicklung in den Dialekten ist kompliziert. Zu diesem Thema siehe Allen 1958, Rix §102,103 und Bartoněk S. 140ff.

Linear B hat zwei Reihen für die Schreibung der Zischlaute, deren Konsonanten mit  $\langle z \rangle$  und  $\langle s \rangle$  transliteriert werden. Welche Laute mit diesen Zeichen bezeichnet wurden ist unklar. Oft wird  $\langle s \rangle$  als einfacher Sibilant und  $\langle z \rangle$  als Affrikate verstanden, jedoch ist es sehr gut möglich, dass es sich tatsächlich um einen Unterschied der Palatalität handelt und beide Zeichenreihen sowohl Sibilanten als auch Affrikaten bezeichnen können. Die Z-Reihe würde dann einen palataleren Laut als die S-Reihe bezeichnen. Mit dieser Annahme lassen sich die mykenischen Schreibungen und die Lautentwicklung zufriedenstellend erklären, wie sich noch zeigen wird.

Dentale Tenuis (aspirata) +  $*j$  im Anlaut ergibt in allen Dialekten  $s$ , das wohl als urgr. gelten darf. Im Inlaut wird hier  $*ts$  angesetzt, das mit altem  $*ts$  zusammenfällt und sich dann dialektal unterschiedlich entwickelt. In Linear B wird für diesen Laut die S-Reihe benutzt, weswegen eine palatale Qualität sehr unwahrscheinlich ist.

Dieselbe Lautkombination an morphologisch transparenter Stelle ergibt einen anderen Laut, der in alphabetischer Zeit mit dem Resultat der Palatalisierung der Velare zusammenfällt, in Linear B aber mit der S-Reihe geschrieben wird. Vielleicht wurde an dieser Stelle das  $*j$  restituert und die Gruppe  $*tsj$  nachmyk. weiterpalatalisiert, so dass sie mit Resultat der palatalisierten Velare zusammenfiel.

$*dj, *gj > *d\acute{z}$ . Da der hier entstehende Laut in Linear B mit der Z-Reihe geschrieben wird, dürfte dieser Laut palatal gewesen sein. Die genaue Phonetik ist offensichtlich dunkel.

Velare Tenuis (aspirata) +  $*j$  scheint eine spätere Palatalisierung als die der Dentale zu sein. Das Ergebnis wird in Linear B ebenfalls mit der Z-Reihe geschrieben, was eine stimmlose Version des obigen Lautes nahelegt, also  $*t\acute{s}$ .

Labial +  $*j$  ergibt  $pt$  in allen Dialekten, muss jedoch eine Vorstufe urgr.  $*p\acute{c}$  gehabt haben, da das Myk. ein eigenes Zeichen für diesen Laut kennt, der erst später auch für ererbtes  $pt$  verwendet wird.

Die Resonanten  $*l, *r, *n$  ( $*m$  in dieser Position schon mit  $*n$  zusammengefallen) werden vor  $*j$  zu palatalisierten Geminaten. Ob hier Anlaut oder Silbengewicht eine

Rolle spielen, ist noch zu überprüfen.

**23**  $*d/ > *//$  Datierung unklar.

**24** Metathese von  $*t$  neben Okklusiven Nach Rix (§106) werden  $*tk$  und  $*tp$  metathetisiert. Für  $*tp > *pt$  gibt er als Beispiel  $*k^{wid-pe} > \tau\acute{\iota}\pi\tau\epsilon$ , das aber wohl besser als  $*k^{wid-k^{we}} > *k^{wik^{wte}} > \tau\acute{\iota}\pi\tau\epsilon$  zu erklären ist. Will man den rätselhaften Anlaut  $pt-$  als Sandhi aus  $*-t\#\#p-$  erklären, so wäre auch eine Metathese neben Labialen anzunehmen. So kann man die allgemeine Regel aufstellen, dass  $t$  + Okklusiv metathetisiert wird.

Datierung unklar.

#### 4.4 Urgriechisch bis mykenische Zeit

In diesem Abschnitt wird die Lautentwicklung bis zur Zeit, in der die mykenische Überlieferung endet, modelliert. Am Anfang dieser Zeit steht die erste dialektale Aufspaltung. Da außer dem recht einheitlichen myk. Dialekt keine weiteren aus dieser Zeit überliefert sind, wird der Lautstand des Myk. – sofern es keinen Grund gibt, etwas anderes anzunehmen – als repräsentativ für das gesamte gr. Sprachgebiet gesehen.

Die dialektale Gliederung ist aus Bartoněk 1991 Tab. B genommen und sieht in Scheme wie folgt aus:

```
(urgr (nwgr (dor nwdor sardor sdor)
        (aiol boiot (thess wthess othess) lesb))
 (sogr (ach myk ark kypr pamph)
        (ion-att ion att euboi)))
```

**25**  $*ti > si$  Dieser Lautwandel wird als grundlegende Unterscheidung zwischen Südost- und Nordwestgriechisch gesehen, ist aber weniger Lautgesetz als einem lieb wäre (Rix §101). Auch westgr. Dialekte haben  $si$  in  $ti$ -Abstrakta und das Myk. geht in  $ko-ri-si-o$  weiter als andere Dialekte, die alle  $\text{Κορίνθιοι}$  zeigen (Bartoněk 2003, S. 144). Hier wird der Lautwandel wegen z.B.  $\epsilon\sigma\tau\iota, \kappa\tau\acute{\iota}\sigma\iota\varsigma$  nach  $s$  und  $k$  ausgesetzt. Eine genauere Lautumgebung – falls man sie überhaupt angeben kann – ist unsicher. Vermutlich ist der Wandel schlicht nicht lautgesetzlich sondern als eine Verallgemeinerung von Sandhivarianten zu sehen (und somit Analogie).

## Literatur

Bartoněk, Antonín (1991). *Grundzüge der altgriechischen mundartlichen Frühgeschichte*. Innsbruck.

- Bartoněk, Antonín (2003). *Handbuch des mykenischen Griechisch*. Heidelberg: Winter.
- Buck, Carl Darling (1910). *Introduction to the study of the Greek dialects: grammar, selected inscriptions, glossary*. Boston: Ginn und Company.
- Rix, Helmut (1992). *Historische Grammatik des Griechischen*. 2., korrigierte Auflage. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Sihler, Andrew L. (1995). *New Comparative Grammar of Greek and Latin*. Oxford/New York: Oxford University Press.