LINGUA POSNANIENSIS
LXVII
2025 (2)

THE POZNAŃ SOCIETY FOR THE ADVANCEMENT OF THE ARTS AND SCIENCES
PHILOLOGICAL AND PHILOSOPHICAL SECTION
THE COMMITTEE OF LINGUISTICS
in co-operation with
ADAM MICKIEWICZ UNIVERSITY IN POZNAŃ

LINGUA
POSNANIENSIS

REVIEW OF GENERAL
AND COMPARATIVE LINGUISTICS

LXVII (2)

POZNAŃ 2025

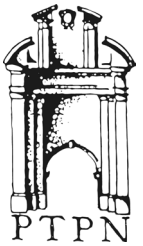
PUBLISHING HOUSE OF THE POZNAŃ SOCIETY FOR THE ADVANCEMENT
OF THE ARTS AND SCIENCES

P O Z N A Ń S K I E T O W A R Z Y S T W O P R Z Y J A C I Ó Ł N A U K
WYDZIAŁ FILOLOGICZNO-FILOZOFICZNY
KOMISJA JĘZYKOZNAWCZA
we współpracy
z UNIWERSYTETEM IM. ADAMA MICKIEWICZA W POZNANIU

LINGUA
POSNANIENSIS

CZASOPISMO POŚWIĘCONE JĘZYKOZNAWSTWU
PORÓWNAWCZEMU I OGÓLNEMU

LXVII (2)



POZNAŃ 2025

WYDAWNICTWO POZNAŃSKIEGO TOWARZYSTWA PRZYJACIÓŁ NAUK

POZNAŃSKIE TOWARZYSTWO PRZYJACIÓŁ NAUK
THE POZNAŃ SOCIETY FOR THE ADVANCEMENT OF THE ARTS AND SCIENCES

GŁÓWNY REDAKTOR WYDAWNICTW PTPN / DIRECTOR OF THE PUBLISHING HOUSE
Jakub Kępiński

KOMITET REDAKCYJNY / EDITORIAL BOARD

Redaktor naczelny / Editor-in-chief
Władysław Zabrocki

Redaktorzy tematyczni / Thematic editors
Leszek Bednarczuk, Nicole Nau, Krzysztof Stroński

Redaktor wykonawczy / Executive editor
Marcin Michalski

Redaktorzy językowi / Language editors
Michael Farris, Stefan Wiertlewski

Sekretarze / Secretaries
Szymon Grzelak, Paweł Kornatowski

Członkowie / Members
Tadeusz Batóg, Andrzej Bogusławski, Vit Bubenik, Eystein Dahl, Armin Hetzer, Henryk Jankowski, Ronald Kim, Tomasz Krzeszowski, Leonid Kulikov, Henrik Liljegren, Terumasa Oshiro, Dennis R. Preston, Gábor Takács, Boris Zakharyin

Poprzedni redaktorzy naczelni / Former Editors-in-chief
Ludwik Zabrocki, Jerzy Bańczerowski

RECENZENCI / REVIEWERS

Mustapha Aabi (Agadir), Carmen Berlinches (Madrid), Paulina Bortnowska (Poznań), Jean Pierre Daquila Ribeiro (Madrid – Barcelona), Jacopo Falchetta (Bergamo), Ignacio Ferrando (Cádiz), Rita Francia (Rome), Sarali Gintsburg (Madrid), Jairo Guerrero (Aix-en-Provence – Marseille), Jan Hoogland (Nijmegen), Amr Keleg (Abu Dhabi), Ronald Kim (Poznań), Sjaak Kroon (Tilburg), Cristina La Rosa (Catania), Ayoub Loutfi (Casablanca), Daniela Merolla (Paris), Francisco Moscoso (Madrid), Marta Noińska (Gdańsk), Nadi Nouari (Seville), Dariusz Piwowarczyk (Kraków), Paweł Rydzewski (Warsaw), Lameen Souag (Villejuif), Krzysztof Stroński (Poznań), Rafał Szeptyński (Kraków), Wojciech Zeyland (Poznań)

Adres Redakcji / Editorial Office
Uniwersytet im. Adama Mickiewicza
al. Niepodległości 4, 61-874 Poznań, Poland
e-mail: linguapo@amu.edu.pl

Informacje dotyczące wskazówek dla autorów są dostępne online pod adresem:
https://pressto.amu.edu.pl/index.php/linpo/

All information regarding notes for contributors is available online at:
https://pressto.amu.edu.pl/index.php/linpo/

Publikacja sfinansowana przez Uniwersytet im. Adama Mickiewicza w Poznaniu.

This publication has been financed by Adam Mickiewicz University in Poznań.

Copyright © by PTPN Poznań 2025

ISSN 0079-4740
eISSN 2083-6090

Wersja drukowana czasopisma jest wersją pierwotną.

TABLE OF CONTENTS

ARTICLES

Timur Akishev, *Loanword adaptation parameters in contact: The case of -ING Anglicisms in Russian* 7

Ahmed Ech-Charfi, *Morphological and semantic opacity as factors of linguistic change: A study of Form VIII verbs in Arabic* 23

Marianna Pozza, S*ynchronic morphological parsing and suffixal productivity in the Hittite adjective: Considerations on some illustrative examples* 41

Wojciech Sowa, *A 'norm' in the Ancient Greek dialectal lexicon? A comparative study of Gregory of Corinth's Περὶ Αἰολίδος and the anonymous Γλῶσσαι κατὰ πόλεις (Urb. Gr. 157)* 65

REVIEW

Alfred F. Majewicz, *Carmen Dagostino, Marianne Mithun, and Karen Rice (eds.) 2023 (vol. 1), 2024 (vol. 2)*. The languages and linguistics of indigenous North America: A comprehensive guide, Volume 1; Volume 2. *(The World of Linguistics 13.1-2). Berlin – Boston: De Gruyter Mouton. Pp. LII (front matter) + 1-715 + separate enlarged map insert (vol. 1), and + pp. XII (front matter) + [987] 717-1702 (vol. 2)* 87

DOI: 10.14746/linpo.2025.67.2.1

Loanword adaptation parameters in contact: The case of -ING Anglicisms in Russian

Timur Akishev

KIMEP University, Almaty
t.akishev@kimep.kz | ORCID: 0009-0009-3684-5500

Abstract: This study focuses on a corpus-based description of the process of linguistic adaptation of nominal English loanwords in Russian containing the deverbal suffix -ING, transliterated into Russian as *инг* /ɪng/. 89 loanword items were analyzed in terms of their linguistic characteristics, such as morphological structure and frequency of occurrence in the corpus. The study discusses the complex nature of the lexical items that are transferred from one language into another through ongoing language contact. The corpus-based analysis included such procedures as the search for any relevant items within a given time frame in a selected corpus, the identification and assigning of language-related characteristics to the items elicited from the corpus, and the statistical procedure that aimed to determine and describe the relationships that exist, or are likely to exist, between different types of characteristics of the loanwords.

Keywords: Anglicisms, Russian, corpus, loanword adaptation, linguistic parameters

1. Introduction

The Russian National Corpus (*Natsional'nyy korpus russkogo yazyka*), henceforth referred to as the RNC, was used as the main source of data for this study. This corpus, abundant in examples of real-life usage of the lexical stock of the Russian language, comprises a number of subcorpora. The current study only considered (1) the main subcorpus, which contains approximately 289 million words; (2) the newspaper subcorpus, which includes some 227 million words; and (3) the spoken subcorpus, nearing 12 million words. Importantly, the aforementioned statistics are provided only as of 2021, which is the year when the searches were conducted. The RNC has since been updated in terms of its content.

The specific time frame that was considered is from 2015 to 2020, representing an update of the body of knowledge on the state of Anglicisms in Russian. Although certain subcorpora of any corpus may not be entirely or equally up-to-date, preliminary corpus searches in the

LP LXVII (2), 2025. © The Author(s). This is an open access article licensed under the CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en/).

RNC revealed a significant number of items with which to work across the subcorpora. The lexical material identified through the preliminary and all subsequent searches differed on an item-by-item basis, which was still useful in constructing a large and diverse database for further analyses. Some of the lexical material was overlapping, with identical or similar items occurring within and across a number of subcorpora of the given corpus. Such a frequent occurrence was still useful in providing more information about the peculiarities of the linguistic behavior of loanword items in the language.

The present study aimed to identify and assign the linguistic characteristics to the loanword material after conducting the first series of searches for any relevant items in the RNC within the above-mentioned time frame. The description and categorization of each identified loanword item relied on a number of systemic parameters: frequential (based on their absolute frequency of occurrence in the RNC) and productivity-based (associated with their capacity for derivational adaptation). Thus, each item in the constructed database had a certain frequency value and a productivity value based on the presence or absence of derivational properties.

Subsequently, a twofold statistical procedure was performed using this database to see how different characteristics of the loanwords correlated with each other. Specifically, the first statistical test focused on the connection between each independent parameter of the loanwords with their frequency of occurrence. In other words, each separate parameter underwent a correlation analysis to establish any significance of its interaction with the frequency characteristic. In this case, frequency was considered as more of an extralinguistic and quantitative property, while productivity was delineated as purely linguistic or language-related parameter. The subsequent statistical analysis, linear regression, determined whether or not the independent linguistic parameter was likely to predict or influence the frequency of occurrence of the loans.

Thus, the main purpose of the current study was to determine and discuss the main peculiarities of linguistic adaptation of the -ING Anglicisms in Russian, with a view of explaining whether and how the loans' morphosemantic characteristics are connected with their frequency of occurrence in this recipient language. In this study, the term "morphosemantic" is used to refer to the aforementioned parameter of productivity.

The main objectives of the current study were as follows:

(1) To determine the characteristics of linguistic integration of -ING Anglicisms into Russian, by means of a comprehensive review of the literature; and
(2) To develop and implement an approach to the statistical analysis of corpus-elicited data on the -ING loans for the period of 2015-2020.

The first objective was addressed by means of a comprehensive discussion of the linguistic nature and behavior of Anglicisms in Russian. Various approaches to defining the notion "Anglicism" were discussed, relying on the existing body of knowledge on English loanwords in Russian represented by the works of American, Russian, Soviet, and international scholars in the fields of Russian-English language contact, bilingualism, and loanword studies. As per the second objective, data were elicited from the RNC using a strict set of procedures that focus on the identification of the respective characteristics of the loanword

material. This step involved a detailed discussion of the parameters and the reasons for which they were selected, the application of the parameters in the process of elicitation of the data from the subcorpora, and the preparation of the database for a subsequent two-step quantitative analysis.

The current study aimed to provide a comprehensive answer to the following two-part research question:

RQ1. *Is there a statistically significant correlation between the parameter of frequency of occurrence of the* -ING *Anglicisms and their productivity-based characteristic? If yes, to what extent can the parameter of frequency be predicted by the parameter of productivity?*

This research question was addressed by means of conducting two distinct statistical procedures, correlation and linear regression. An initial database was developed based on the corpus searches, including the inherent frequency characteristic of each item. This database was further modified in terms of the application of the parameter-driven approach.

Notably, the current study focused only on analyzing the productivity-frequency relationship. In a pilot study (Akishev 2023), the quantitative analyses into the parameters of structure and meaning did not provide any statistically significant results in relation to the parameter of frequency, but they were still important in (1) describing the linguistic nature of the Anglicisms, (2) discussing the distribution of different morphological and semantic types of the Anglicisms in terms of their frequency in the corpus, and (3) analyzing the relationships between the morphosemantic nature of the loanwords and their capacity for derivational productivity.

In the pilot study, the analyses of the parameter of structure did not yield any statistical significance, but the construction of the working database still revealed a number of interesting patterns (Akishev 2023: 42). The majority of the -ING loans in the corpus for the given period were of the simple, monolexemic structural type. It was also found that the simple-type loans have a higher frequency of occurrence in the corpus. The morphological nature of these cross-linguistic constructions is essential to any discussion of the regular patterns of their integration into the recipient language. However, a number of irregularities in the structural nature of some of the Anglicisms were identified. For example, the Anglicism *bukkrossing* ‘bookcrossing’ comes from the bilexemic compound ‘Book’ + ‘Crossing’ in which the first letter of each lexeme is capitalized, likely due to the fact that this compound originated in English from the name of a brand and a website that were subsequently popularized. In Russian, however, this capitalization was not preserved when the word was borrowed and adapted. Another interesting example is the Anglicism *smoking*, which comes from the nominal construction ‘smoking jacket.’ In the process of adaptation into Russian, the second part of this construction, ‘jacket,’ was omitted, perhaps for the purpose of simplification – shortening the loanword to make it more easily adapted and widely used, or in order to make this loan structurally resemble many other loans ending in -ING that do not usually have another lexeme attached after the main -ING lexeme. Finally, the construction *kaming-aut* ‘coming-out’ is another interesting example. In this example, the second part of the loan is preserved, although it is not a second lexeme, as in ‘smoking jacket.’ The second part ‘out’ in ‘coming-out’ is a particle that is connected in a phraseological manner to the

-ING lexeme, which is why the omission of the second part of the loan would be detrimental to the preservation of the semantic integrity of this construction.

Pilot statistical analyses of the characteristics of the loans' meaning also provided no numerically significant results. However, the descriptive statistics analysis demonstrated that the majority of the -ING Anglicisms in the corpus pertain to the core semantic type (Akishev 2023: 44). This type of Anglicisms refers to notions that are already familiar to the language and culture, while the cultural type is used to denote unfamiliar and new concepts (Calude et al. 2020). A large number of the -ING loans of the cultural meaning type are related to sports: *snoubording* ('snowboarding'), *skeitbording* ('skateboarding'), *sapserfing* ('SUP surfing'), *veikbording* ('wakeboarding'), *kikboksing* ('kickboxing'), *kerling* ('curling'), *vindserfing* ('windsurfing'), *bouling* ('bowling'), *chirliding* ('cheerleading'), and *serfing* ('surfing'). These names entered Russian culture via -ING Anglicisms as absolutely new kinds of sports that previously did not exist for the Russian-speaking society. The remaining cultural loans are related to such semantic fields as technology, beauty, economics, and lifestyles.

The scrutiny of four parameters in the pilot study provided certain insights into the specificities of the adaptation of Anglicisms in Russian. Building on those findings, the current study focuses more narrowly on the productivity-frequency relationship. The main concepts that are presented above are discussed in detail in the following section, including, but not limited to, the notion "Anglicism," the integration of Anglicisms in Russian, and the parameters selected to describe the linguistic characteristics of the loanwords.

2. Previous research on Anglicisms in Russian

Attracting attention due to their popularity, role and functions, and complex cross-linguistic nature, Anglicisms have long been the subject of scrutiny in Russian linguistics. According to the most commonly used definition, an Anglicism is a word or phrase of English origin which is borrowed into another language. Dyakov (2012: 73) defined the notion "Anglicism" as a unit of a certain linguistic level that was transferred into Russian and underwent a certain alteration. While this definition emphasizes the processes of transference and transformation, it does not mention the role and functions or the contribution that Anglicisms make. On the other hand, these aspects were already discussed in detail by scholars like Benson (1959), who posited that the vast majority of Anglicisms are nouns. In certain cases, these nouns were transformed into adjectives by means of derivation on the basis of the rules of Russian morphology. Benson (1959: 257) singled out the following four groups of Anglicisms based on their degree of assimilation into Russian: Group 1 includes loanwords that have become an inherent part of active Russian vocabulary (e.g., *film*, *klub*); Group 2 includes mainly scientific terms which are not used commonly by all speakers (e.g., *biheviorizm*); Group 3 comprises loanwords which describe non-Russian cultural patterns (e.g., *kolledzh, mister*); and Group 4 includes obsolescent words that were borrowed but not subsequently picked up by speakers, thus losing in the competition with their native Russian counterparts (e.g., *breikwoter – volnolom*). An important aspect of Benson's approach was the emphasis on the inherent "nounness" of Anglicisms. Furthermore, he stressed the fact

that some Anglicisms can undergo certain morphological transformations on the basis of Russian morphology. However, the reasons for such changes and the process of borrowing as such were not discussed.

A more recent approach by Wójtowicz (1995) filled in those gaps through a comprehensive discussion of Anglicisms in Russian. The author also argued that the majority of these loanwords are nouns. As for the adaptational changes that they undergo, Anglicisms generally acquire the grammatical category of masculine gender in Russian (e.g., *klub*, *kolledzh*), which is an open-class system. They appear in Russian as a result of direct contact between the two languages, with no traces of interference from other languages. This perspective incorporates the previously discussed theoretical concepts underlying Anglicisms and develops them further, stressing that the main conduit of noun loans from English into Russian is direct and uninterrupted language contact.

More specific reasons for the borrowing, adaptation, and use of Anglicisms were provided by Styblo (2007: 64): (1) the loanword is characterized by novelty and positivity of its connotational parameters; (2) the original Russian word is outdated and has negative connotations; and (3) the loanword can be more suitable to express the speaker's purport. Unlike the previously discussed approaches, Styblo's point of view emphasized the semantic aspects of the phenomenon under study. This perspective would be useful in expanding other approaches, such as, for example, Dyakov's (2012) previously discussed definition of the notion "Anglicism." However, while Styblo's approach focuses on the semantics of Anglicisms, it still omits the pragmatic aspects, especially the speaker's attitudes and identities involved in the process of adaptation and use of Anglicisms.

Traditionally, scholarly interest has centered on the grammatical properties that Anglicisms acquire upon integration into Russian (Aristova 1978; Vorobyova 2009), along with their semantic and stylistic characteristics and the pragmatic functions that they perform. More recent research has examined the role of Anglicisms in shaping discourse (Privalova et al. 2024), their linguistic features in contemporary slang (Lacková 2021), their contribution to youth culture (Eshenkova 2023), and their functions in diaspora dialects of Russian (Makarova 2022).

What has not gained enough attention in the research are the peculiarities of relationships across different linguistic features of the loans. One such feature, whose relation to the other characteristics of Anglicisms deserves more attention, is morphological productivity. This phenomenon can be defined as the capacity to develop into more linguistically complex structures based on the principle of morphological derivation (Booij 2010: 8).

As evidenced by the richness and flexibility of Russian word formation, the -ING Anglicisms often tend to transform into more complex linguistic structures. Apart from assimilation into Russian as full-fledged nouns (Vorobyova 2009: 183), these loans can also be morphologically transformed into verbs. Dyakov (2001: 161) argues that their transformation into verbal constructions is based on the morphological process of suffixational derivation characteristic of solely the Russian language. Discussing the morpheme -ING which is transferred together with the noun loans ending in it, Dyakov concludes that -ING is a productive morpheme, which is evidenced by its ability to construct new Russian words based on linguistic creativity and pertaining to the spoken vocabulary (2001: 201): for example,

shippingovat' 'to support a romantic relationship between two or more fictional characters or real-life people.' It can be deduced from Dyakov's standpoint that the adaptation of Anglicisms in Russian is a multifaceted phenomenon, as well as a process that may involve multiple linguistic mechanisms pertaining to the levels of phonology, morphology, syntax, semantics, and pragmatics.

Prokutina (2012: 8) argues that Anglicisms and their morphological characteristics are primarily influenced by the Russian derivational system. She also claims that the main criterion of adaptedness of Anglicisms is their ability to interact with original word formation means of Russian. According to the author, -OVAT' is one of the most productive suffix constructions used in the process of derivation of verbs from noun loans. Importantly, the current study also regards this suffix construction as one of a number of Russian suffixes that may interact with the -ING loans, not the only suffix to be able to do so. The main focus, however, is on the morphological transformation of the loans from nouns into verbs. The analysis of the peculiarities of such transformations serves as a foundation for the discussion of the derivation-based productivity of the loans as one of a number of their linguistic parameters.

Previous corpus-based studies have also informed the approach and the choice of parameters of adaptation of the loanwords. These works were selected due to their foci on different aspects of loanword adaptation; namely morphology, semantics, and quantitative characteristics. The approach to the quantitative features of the Anglicisms that was implemented in this study is corpus-based. It was informed by the methodological framework developed by Vardøy & Post (2013), who conducted a corpus-based analysis of -ING loanwords in Russian dated between 2000 and 2011 in the newspaper subcorpus of the RNC. Based on data from the corpus, the authors compiled a list of -ING Anglicisms and their frequency of occurrence, and commented on the loanwords' popularity over time. Very interesting is the authors' decision to restrict the compilation of their database to the newspaper part of the RNC, which they considered to be the most representative of the Russian language in action. The current study will include two more parts of the RNC in the analysis. Furthermore, Vardøy and Post focused primarily on the semantics of these loanwords, which, as they said, is an underresearched area in loanword studies. This study replicates some aspects of these authors' research and updates it for the specific period of 2015-2020. In addition, apart from considering both semantic and quantitative characteristics of these loans, due attention will also be paid to the aspects of their morphological adaptation and productivity in the recipient language, which is why the present study is also an extension of Vardøy & Post's approach.

The parameters of productivity and frequency of occurrence will play a pivotal role in the course of the statistical analyses aimed at providing a comprehensive answer to the research question. Every item drawn from the RNC for the period stated will possess a certain value according to this parameter. A dichotomous approach is of great importance to the analysis of the loans' characteristics. Thus, the parameter of productivity will be analyzed using a binary system (productive vs. non-productive). Each -ING loanword item drawn from the RNC will possess this characteristic in addition to the frequency of its occurrence in the corpus.

3. Research design

3.1. Corpus elicitation approach

The database will be compiled by means of eliciting data from the RNC from 2015 to 2020. Searches for words ending in -ING will be performed in its three different subcorpora, with each of the searches modified similarly in order to obtain data which are as much as possible representative of the actual situation in respect of loanword adaptation in Russian. Data will be drawn from three distinct subcorpora: main (289 million words), newspaper (227 million words), and spoken (12 million words). The decision to choose these three different subcorpora, as opposed to working simply with one of them, as Vardøy and Post (2013) did with the newspaper subcorpus, is based on the fact that they present examples from three different dimensions of language use, which definitely increases the amount of data a researcher can work with and can provide them with more diverse, relevant, and novel language material.

The RNC is up to date, with over 600 million instances of use of the Russian language in multifarious contexts. Regular updates incorporate more language material into multiple subcorpora. According to the website, the RNC's main purpose is to serve as a useful tool for anyone interested in Russian (such as professional linguists, language teachers, and language learners). Two versions of the RNC are currently available: the old and the new. The new version is set to replace the old one because it is more technologically advanced, with a number of search mechanisms improved. The new version has a more user-friendly interface and a more straightforward search system, which is why it was used for the purposes of this study.

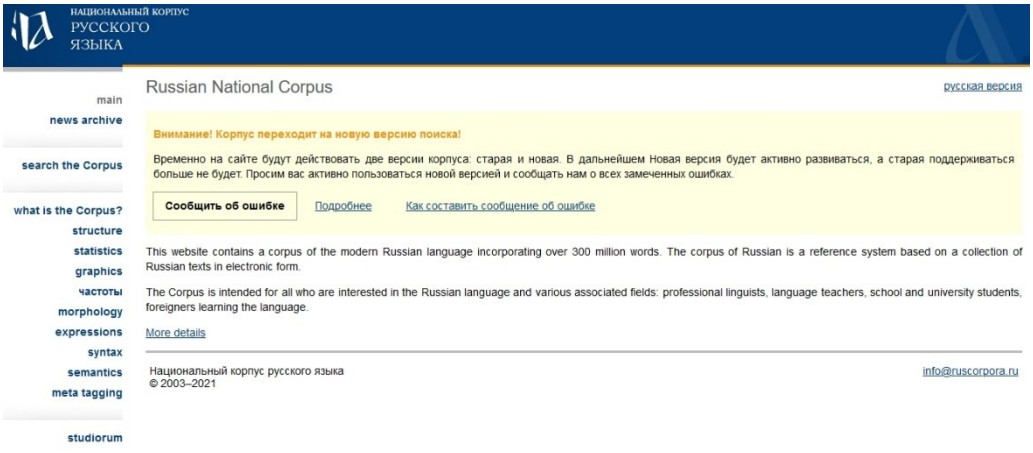


Figure 1: The website of the Russian National Corpus. *Natsional'nyy korpus russkogo yazyka* ©2021. (https://ruscorpora.ru/new/en/) (Accessed 2021-04-07)

According to Lyashevskaya et al. (2003: 113), the Russian National Corpus (https://ruscorpora.ru/en) search system can be used in two formats: disambiguated and non-disambiguated. The non-disambiguated format allows units possessing the same orthographic structure to preserve their ambiguity independent of the context in which they occur. The disambiguated format does exactly the opposite. In the search system located on the website of the RNC, users can limit their searches either to the disambiguated or non--disambiguated corpus search parameter. The use of the non-disambiguated parameter results in a much larger number of items found in the RNC due to the fact that homonymous units acquire all possible sets of morphological characteristics (Lyashevskaya et al. 2003: 114). This study will use the non-disambiguated setting, so that the results will be richer. Furthermore, unlike Vardøy & Post (2013), capitalized words will not be deleted from the results. Although this action does indeed narrow the search by omitting proper names, it also eliminates a substantial number of clause-initial items, which may result in a misinterpretation of the actual situation with the number of loanwords.

3.2. Parameter-driven approach

The first parameter of the loans is frequency. It is different from the other parameter in that it is assigned automatically based on corpus-provided information, which is why this parameter is considered to be inherent to the data. The parameter of frequency will be based on the total number of examples that a given loanword has in each of the three subcorpora. The RNC provides a table at the bottom of the search results page which contains the list of items found and the total number of examples for those items across the sum total of search results for a given period of time. Based on results from all three subcorpora, a single database will be compiled, in which the items will be ordered and any irrelevant elements deleted. For the words that are present across the two or three parts of the RNC, the total number will be added up from the numbers determined in the correspondent parts. The words which do not overlap will be entered into the database in descending order of numeric magnitude. The final database will include the loans and their respective total numbers of occurrences not per each of the subcorpora, but for all of them combined.

The other parameter will be operationalized in terms of a specific approach to a multiaspectual description and statistical analysis of the corpus data, so it is referred to as a descriptive parameter. This specific parameter-driven approach is aimed at determining whether the inherent parameter, frequency, can be associated with or dependent on, the descriptive parameter.

The productivity parameter is based on the derivational capacity of the -ING loans to interact with the Russian verb-forming suffix -OVAT'. Based on this parameter, productive and non-productive Anglicisms will be identified. The productive type is characterized by the ability to undergo the process of derivation to form Russian verbs using Russian verb morphology (e.g., *shering* 'sharing' + *ovat'* → *sheringovat'* 'to perform sharing/to share'). The non-productive type of Anglicisms, on the other hand, does not possess such a capacity. To determine whether an -ING loan is productive or non-productive, a two-step procedure will be performed. First, corpus searches will be conducted for each item in

combination with the suffix -OVAT' to see if the RNC, as a most credible and up-to-date source of information on the Russian language in use, has any examples of such morphologically complex constructions. Secondly, searches will be conducted for the corresponding word-formation patterns on the basis of two dictionaries: Dyakov (2014) and Yegorova (2014).

3.3 Statistical procedures

3.3.1. Overview

The research question will be answered by means of conducting Point-Biserial Correlation and Simple Linear Regression. The only numeric parameter of the -ING Anglicisms that is discussed in the current study, the parameter of frequency, will serve as the variable against which the other parameter will be analyzed. The purpose of testing against frequency is to see whether it may be associated in a certain way with the descriptive dichotomous parameter. For instance, the descriptive parameter, based on the results of the statistical tests to be conducted, may be significantly correlated with or even predict the frequency of occurrence of the -ING loans. Special attention in this study will be paid to the discussion of the productivity parameter. The total number of items that will be productive based on the database to be compiled using the RNC search results for the stated period will show the current situation with the productive capacity of the -ING Anglicisms.

3.3.2. Point-Biserial Correlation analyses

The procedure incorporates the correlation analysis. Importantly, this test does not distinguish between dependent or independent variables, as both variables are simply correlated. The question that the correlation analysis aims to answer is the following:

(1) Is there a statistically significant relationship between the productivity parameter of -ING Anglicisms and the parameter of frequency?

3.3.3. Simple Linear Regression analyses

The statistical procedure focuses on linear regression. Importantly, frequency is a dependent variable, as the procedure is aimed at testing whether frequency can be affected or predicted by productivity. The question that the given statistical test is designed to answer is as follows:

(1) To what extent can the parameter of frequency be predicted by the parameter of productivity?

4. Results and discussion

4.1. Corpus searches and database compilation

The searches were conducted in the three subcorpora for the period 2015-2020. The lexico-grammatical search option was used to look for any words that have an -ING (Russian инг) ending, using the *инг mask, as the lexeme search option only allows to search for entire words. The first unspecified search in the main subcorpus resulted in a large number of items that had the ending -ING: 46,844 context-based instances in 14,029 documents. Subsequently, the search settings were introduced, such as a non-disambiguated search, both nonfiction and fiction texts, and the 2015-2020 time period. As a result of these modifications of the search settings, a user-generated subcorpus was created with a volume of 731 documents and 2,367,342 words. The second search was conducted for the -ING items in this user-generated subcorpus, which resulted in 484 context-based instances in 182 documents. All items were automatically assigned a certain frequency value by the system. The identified items along with their frequency values were placed in a rough-draft version of the database to be compiled upon gathering all the data from all three subcorpora.

The first undefined search for the -ING items in the newspaper subcorpus of the RNC resulted in 215,426 context-based instances in 107,330 documents. The newspaper subcorpus consists of two parts: *Media of the 2000s* and *The Corpus of regional and foreign press*. As the former is obviously outdated, the search for the Anglicisms was conducted in the latter. The search settings were then applied, which included a non-disambiguated search, nonfiction and fiction texts, creation date between 2015 and 2020, country (Russia), and language of origin (Russian). A user-generated subcorpus was subsequently developed with a volume of 1,085 documents and 567,354 words. The search for -ING items resulted in 329 contexts in 185 documents. Once again, the items elicited were entered into a preliminary database.

The first search in the spoken subcorpus provided 2,177 contexts in 586 documents. Following the implementation of the settings, such as non-disambiguated search and 2015-2020 creation date, the generated subcorpus included 143 documents with a total volume of 371,015 words. Finally, the search in this subcorpus resulted in 66 context-based uses in 30 documents. The items were subsequently copied into the above-mentioned database.

After conducting all the searches in the three subcorpora and obtaining the data, all irrelevant items were deleted. These items included non-Anglicisms (e.g., *viking* ‘Viking’), Anglicisms whose roots include the same letters as the morpheme -ING (e.g., *king* ‘king,’ *ring* ‘ring’), last names ending in -ING (e.g., *Kipling* ‘Kipling,’ *Rouling* ‘Rowling,’ *Turing* ‘Turing,’ etc.), and cross-linguistic multilexemic English-Russian blends (e.g., *vetromonitoring* ‘wind monitoring’). The process of compounding of the -ING Anglicisms with Russian words to produce cross-linguistic blended or conjoined constructions is not relevant to this study. These constructions cannot be easily categorized. Cross-linguistic compounds are too complex morphologically and semantically to be analyzed in terms of the current approach that focuses on English-only monolexemic or bilexemic words. For example, the blended word *rosbizneskonsalting* incorporates three different lexemes:

(1) an original Russian contraction *ros-* of the adjective *rossiyskiy* 'Russian'; (2) a non-ING nominal Anglicism *biznes* 'business' that is quite an old loanword; and (3) an -ING Anglicism *konsalting* 'consulting.'

Subsequently, an initial unparametrized database was developed, in which each of 89 items on the final list of Anglicisms possessed a certain frequency value. The number of occurrences of any identical items that were identified in different subcorpora was added up to reflect the sum total of occurrences across the given corpus. The items were listed in a descending order based on the magnitude of their frequency values. An excerpt from this database including the first ten items on the list is given below (Table 1).

Table 1. Excerpt from the initial unparametrized Anglicism database

#	-ING Anglicism (including the original English words)	Frequency of occurrence in the RNC (total number of items encountered)
1.	*reiting* 'rating'	180
2.	*miting* 'meeting'	59
3.	*monitoring* 'monitoring'	59
4.	*doping* 'doping'	38
5.	*holding* 'holding'	36
6.	*trening* 'training'	30
7.	*marketing* 'marketing'	26
8.	*kasting* 'casting'	17
9.	*brifing* 'briefing'	12
10.	*kemping* 'camping'	12

4.2. Productivity parameter analyses

The goals of the analyses in terms of the parameter of productivity are:

(1) To categorize the items based on the productive/non-productive distinction, using RNC's search system and the previously used two dictionaries of foreign words and Anglicisms;
(2) To conduct a Point-Biserial Correlation with frequency as a continuous variable, and productivity as a dichotomous categorical variable;
(3) To conduct a Simple Linear Regression with frequency as a continuous *dependent* variable, and productivity as a dichotomous categorical *predictor* variable; and
(4) To interpret and summarize the results.

To categorize the items based on their capacity for derivation-based productivity, searches were conducted for every item on the list in the two dictionaries and the RNC in

combination with the verb-forming suffix -OVAT'. The items that were found to exist in the resources in terms of such a morphological construction were entered into the database as productive Anglicisms. The items that were not found to form verbs using that suffix were identified as non-productive Anglicisms. An excerpt from the working database is presented in Table 2. For productive Anglicisms, the resulting Russian verb is presented in this database along with the transliterated components from which it was derived, the -ING Anglicism itself and the suffix OVAT'.

Table 2. Excerpt from the working database for productivity parameter analyses

#	-ING Anglicism (including the original English words)	Frequency of occurrence in the RNC (total number of items encountered)	Derivation-based productivity (productive or non-productive), including the resulting verb and its transliterated structure
1.	*trening* 'training'	30	Non-Productive
2.	*marketing* 'marketing'	26	Non-Productive
3.	*kasting* 'casting'	17	Productive (kastingovat': *kasting* + *ovat'*)
4.	*brifing* 'briefing'	12	Productive (brifingovat': *brifing* + *ovat'*)
5.	*kemping* 'camping'	12	Productive (kempingovat': *kemping* + *ovat'*)
6.	*smoking* 'smoking jacket'	12	Non-Productive
7.	*kraudfanding* 'crowdfunding'	11	Non-Productive
8.	*bukkrossing* 'bookcrossing'	8	Non-Productive
9.	*lizing* 'leasing'	7	Non-Productive
10.	*pressing* 'pressing'	7	Productive (pressingovat': *pressing* + *ovat'*)

A cross-tabular analysis demonstrated the difference in numbers of items within each type of the loans' productiveness. Out of the total list of 89 Anglicisms, 64 items were found to be non-productive and the remaining 25 items were found to be productive. The statistical procedures which follow will elucidate the peculiarities of the relationships of each type of productiveness to the loans' frequency of occurrence in the given corpus.

In terms of the Point-Biserial Correlation test, the frequency variable was considered in relation to the productivity variable. The variances were also found to be unequal, but this fact did not affect the statistical procedure. As a result, a statistically significant correlation-based link was found between the loans' productivity and their frequency of occurrence, $r = .36$, $p = .0001$. The following table illustrates the specifics of this significant relationship.

Table 3. Point-Biserial Correlation results for the frequency-productivity relationship

Correlations			
		Productivity	Frequency_SQR
Productivity	Pearson Correlation	1	.362**
	Sig. (2-tailed)		.000
	N	89	89
Frequency_SQR	Pearson Correlation	.362**	1
	Sig. (2-tailed)	.000	
	N	89	89

**The correlation was significant at the 0.01 level, two-tailed.

The following figure illustrates the positive significant trend observed in terms of the correlation of the parameters. The parameter of productivity is on the x-axis, with 0 signifying the non-productive type and 1 signifying the productive one. The box plot shows that there is a slightly higher frequency of occurrence for productive loans than for non-productive ones. This finding corroborates the assumption that the loans' adaptedness in the recipient language is to a certain extent dependent on their frequency of occurrence in the corpus.

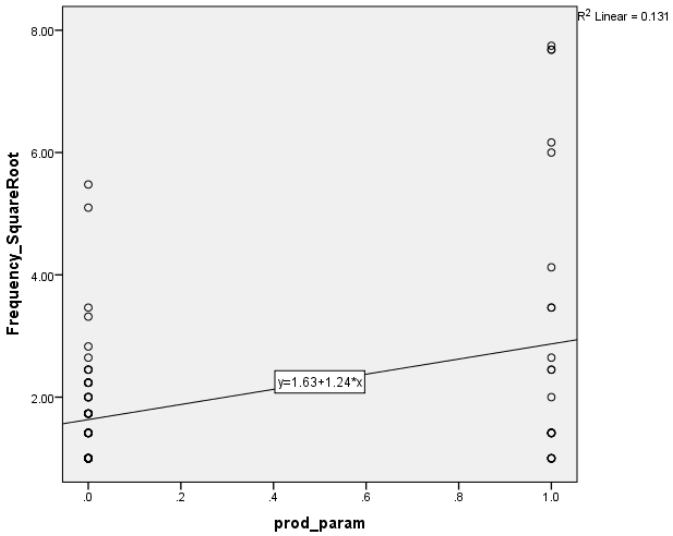


Figure 2: Box plot illustrating the correlation between the parameters of productivity and frequency

These significant findings will play a pivotal role in the forthcoming discussion of the loans' linguistic behavior in Russian. The parameter of productivity will be considered as one of the main factors affecting the loans' frequency of occurrence. The following statistical procedure will determine if the loans' frequency can be predicted by their derivational capacity.

The results of the linear regression analysis demonstrate that there is a statistically significant relationship between the predictor and the outcome variable, $r\,(87) = .362$, $p < .001$. The r^2 for the equation was .131, which means that 13% of the variability in the frequency of occurrence can be explained by the parameter of productivity.

Table 4. Model summary for the linear regression test within the parameter of productivity

MODEL SUMMARY								
				Change Statistics				
Model	R	R Square	Adjusted R Square	R Square Change	F Change	df1	df2	Sig. F Change
1	.362[a]	.131	.121	.131	13.141	1	87	.000

a. Predictors: (Constant), Productivity_Parameter
b. Dependent Variable: Frequency_SquareRoot

In sum, the parameter of productivity was found to be the main parameter that moderately correlates with and predicts the frequency of occurrence of the -ING Anglicisms in the corpus.

5. Conclusion

It was hypothesized that there is statistical significance between the frequency and the descriptive parameters of the -ING loans. The quantitative methodological approach was expected to demonstrate significance of findings both in correlation and regression tests. In terms of the correlation tests, it was expected to obtain findings that would corroborate certain relationships existing between or across parameters. The regression procedure was expected to demonstrate whether or not the descriptive parameter can predict or affect the inherent parameter of frequency.

Both statistical analyses have provided different results: both statistically significant and non-significant. The significance of the findings was determined only as a result of the correlation and regression tests for the parameter of productivity in relation to the Anglicisms' frequency of occurrence in the corpus. However, the scrutiny of all the parameters has provided certain insights into the specificities of the adaptation of Anglicisms in Russian.

The analyses for the parameter of productivity yielded significant results. It was observed that the loans' productivity is moderately correlated with their frequency, which it also may predict to a certain extent, as the regression test demonstrated. Thus, the Anglicisms' productivity type can be characterized by possessing a higher frequency of occurrence in the corpus. In the working database, on the other hand, most of the Anglicisms were of the non-productive type. These Anglicisms were found to possess a lower frequency of occurrence in the corpus. Both statistical procedures have shown that the loans' frequency can be expected to increase if they fall into the productive category, and to decrease if they pertain to the non-productive group. The numerically significant correlation-based interaction between

the parameters of frequency and productivity demonstrates that while the loans' frequency of occurrence can be considered to be the main and most obvious factor of their adaptedness, their productivity characteristic may be an underlying phenomenon affecting and predicting the former. However, it should be understood that the relatively small dataset analyzed in this study constitutes a limitation that restricts the statistical power of the results. Another limitation is connected with its primary focus on the morphological and productivity-related aspects of the loanwords, while a more detailed discussion of the semantic and pragmatic characteristics would provide a fuller picture of the phenomenon.

The findings obtained indicate that there is a certain association between the descriptive parameter of derivation-based productivity and the inherent parameter of frequency. It is hoped that this study will be conducive to the expansion of the body of knowledge on Russian--English language contact, Anglicisms in Russian, loanword morphology and semantics, loanword adaptation, and replicable in corpus linguistics research and quantitative methods in applied linguistics.

References

Akishev, T. 2023. *Morphosemantic integration of* -ING *Anglicisms into Russian and Kazakh in the context of trilingual code-switching.* Oxford: The University of Mississippi. (Doctoral dissertation). (https://egrove.olemiss.edu/etd/2656) (Accessed 2024-11-13.)

Aristova, V. M. 1978. Morfologicheskaya struktura angliyskikh slov v russkom yazyke. In Aristova, V. M., *Anglo--russkiye yazykovyye kontakty*, 44-47. Leningrad: Izdatel'stvo Leningradskogo Universiteta.

Benson, M. 1959. English loanwords in Russian. *Slavic and East European Journal* 3. 248-267.

Booij, G. 2010. Construction morphology. *Language and Linguistics Compass* 4(7). 543-555.

Calude, A. S. & Miller, S. & Pagel, M. 2020. Modelling loanword success: A sociolinguistic quantitative study of Māori loanwords in New Zealand English. *Corpus Linguistics and Linguistic Theory* 16(1). 29-66.

Dyakov, A. I. 2001. *Derivatsionnaya integratsiya anglitsizmov v russkom iazyke kontsa XX-go veka v funktsional'nom aspekte*. Novosibirsk: Novosibirsk State Pedagogical University. (Doctoral dissertation.)

Dyakov, A. I. 2012. Anglitsizmy: Zaimstvovanie ili slovoobrazovanie. *Filologicheskiye nauki* 5. 72-76.

Dyakov, A. I. 2014. *Dictionary of Anglicisms of the Russian Language*. (http://anglicismdictionary.dishman.ru/Slovar) (Accessed 2024-12-01.)

Eshenkova, N. L. 2023. Anglitsizmy kak chast' sovremennoy molodezhnoy kultury. *Russian Linguistic Bulletin* 9(45). 1-5.

Lacková, M. 2021. Morphological peculiarities of lexical units of English origin in contemporary Russian slang: Dictionary and corpus analyses. *RUDN Journal of Language Studies, Semiotics and Semantics* 12(3). 632-651.

Lyashevskaya, O. N. & Plungyan, V. A. & Sichinava, D. V. 2003. O morfologicheskom standarte Natsional'nogo korpusa russkogo iazyka. *Natsional'nyy Korpus Russkogo Iazyka* 2005(2). 111-135.

Makarova, V. 2022. Anglicismes en russe doukhobor. *Espaces Linguistiques* 4. (https://www.unilim.fr/espaces-linguistiques/478) (Accessed 2025-12-28.)

Privalova, I. V. & Petrova, A. A. & Gishkaeva, L. N. 2024. Metody monitoringa anglitsizmov v russkoyazychnom molodezhnom diskurse. *Russian Language Studies* 22(4). 633-661.

Prokutina, Y. V. 2012. Angliyskiye zaimstvovaniya v nestandartnoy leksike sovremennogo russkogo yazyka: strukturno-slovoobrazovatel'nyy aspekt. Tobolsk: Tobolsk State Social and Pedagogical Academy named after D. I. Mendeleev. (Extended abstract of Doctoral dissertation.)

Russian National Corpus – http://ruscorpora.ru/en/search-main.html

Styblo Jr, M. 2007. *English loanwords in modern Russian language*. Chapel Hill: The University of North Carolina at Chapel Hill. (Doctoral dissertation).

Vardøy, B. F. & Post, M. 2013. A corpus-based study of English loanwords on *ing* in Russian (2000-2011). *Bergen Language and Linguistics Studies*. 3(1). 179-194.

Vorobyova, S. V. 2009. Grammaticheskaya assimiliatsiya noveyshikh anglitsizmov v russkom yazyke. *Bulletin of the Minsk State Linguistic University* 5(42). 178-186.

Wójtowicz, M. 1995. O novykh angliyskikh zaimstvovaniyakh v sovremennom russkom yazyke. *Studia Rossica Posnaniensia* 26. 137-145.

Yegorova, T. V. 2014. *Slovar' inostrannykh slov sovremennogo russkogo yazyka*. Moscow: Adelant.

DOI: 10.14746/linpo.2025.67.2.2

Morphological and semantic opacity as factors of linguistic change: A study of Form VIII verbs in Arabic*

Ahmed Ech-Charfi

Mohammed V University in Rabat
a.echcharfi@um5r.ac.ma | ORCID: 0000-0002-4946-1380

Abstract: This paper deals with the role of morphology in the reconstruction of lexical meaning. It focuses on the case of the Arabic Form VIII verbs in order to illustrate the challenge that morpho-semantics presents for historical lexicographers assuming the unity of a language throughout a long period of its use. In this connection, the paper attempts to show that, although Form VIII verbs have been in use since the early stages of Arabic, it is likely that users assigned them different meanings according to whether Form VIII morpho-semantics was transparent or opaque. Three factors have been identified that increase the opacity of this category: allomorphy, polysemy and frequency of the derivation base. 529 items were culled from a bilingual dictionary for the purposes of the study, and allomorphy was found to contribute about 12% to morphological opacity, and more than 70% of the verbs had a non-prototypical sense. Many of the extended senses seem to have lost all kinds of semantic relation to the prototypical sense, thus resulting in less transparency in the semantics of the derived forms. The study also argues that the less frequent the base of the derivation is, the more opaque Form VIII will be. The paper concludes that, given the lack of rich data from the early stages of Arabic, it is likely that a satisfactory reconstruction of the meaning of derived forms will probably never be achieved.

Keywords: morphology, semantic change, word meaning, Arabic, Form VIII verbs

1. Introduction

Reconstructing lexical meanings of morphologically complex words can be particularly challenging and rather imprecise. The challenge, as will be argued in this paper, is posed mainly by the intricate polysemy networks that derivational morphemes develop over time as well as the phonological changes these morphemes undergo. When these two factors are combined, the relation between meaning and form can become extremely opaque so much

* Some changes have been made to an earlier version of this paper, including the addition of some references, as a response to remarks made by an anonymous reviewer, which are thankfully acknowledged.

LP LXVII (2), 2025. © The Author(s). This is an open access article licensed under the CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en/).

so that native speakers fail to recover it from use, a fact that calls for reanalysis. In the case of languages like written Arabic, which has been in use for around a millennium and a half, it is a risky mission to determine whether and when such a reanalysis happened and what its consequences on the grammar of the language were. The study will focus on Arabic verbal morphology, as represented by Form VIII verbs.

Arabic verbal morphology is very elaborate. It relies on affixes (essentially prefixes and infixes) to derive new verb forms with various meanings. For example, the augmented form *ʔafʕal* is derived from the basic form *faʕal* to convert an intransitive verb into a transitive one, e.g. *xarağ* 'to get out' vs. *ʔaxrağ* 'to put s.o/s.th out'. Similarly, *ftaʕal* is augmented with a *-t-* infix that expresses the notion of reflexivity and related meanings, e.g. *bāʕ* 'to sell' vs. *btāʕ* 'to buy'. As is often noted, derivational morphemes tend to be restricted in productivity but highly polysemous (Bauer 2001, Kotowski & Plag 2023). The infix *-t-*, for example, is limited to a small set of verbs, e.g. **ʕtamal,* from *ʕamal* 'to do', is not attested, though conceptually possible with the meaning 'to do s.th for oneself'. In turn, the infix can express a variety of meanings, besides that of reflexivity, as will be explained later. In diachronic studies on Arabic, the semantic content of morphological forms tends to be overlooked, probably because of their polysemous nature that often escapes satisfactory characterization. This problem becomes even more acute when there is no community of native speakers to whom researchers can turn for nuanced judgments about meaning and use. For example, the online Doha Historical Dictionary of the Arabic Language (*Muʕğam*) provides the following allegedly original meaning of the verb *ḥtaram:* "*waqqara-hu wa rāʕā mā yağibu min ḥaqqi-hi*" ('to respect s.o and to acknowledge his due'). This definition, however, seems to reflect modern use rather than the old use that goes as far back as pre-Islamic times, as this study will show.

The most problematic aspect of this reconstructed meaning is the total disregard of the morphological structure and its contribution to the sense of the verb. As it stands, *ḥtaram* is an augmented Form VIII verb that is derived by the insertion of the affix *-t-* after the first root consonant of its Form I equivalent. The basic form from which it is derived should be *ḥaram* (to deny s.o s.th) or *ḥarum* (to be forbidden), assuming that the *t* after the first consonant is an infix. Given the reflexive meaning of the affix *-t-* to be discussed later, *ḥtaram* should mean something like 'to deny oneself s.th' . The root 'Ḥ-R-M' also refers to the sanctity of the thing forbidden, probably because of its association with the Kaaba called *al-bayt al-ḥarām* (literally, the forbidden house) and, later on, with the Islamic notion of *ḥarām* (sin, wrongdoing, etc.). This shade of meaning can be detected in other related words like *ḥaram* (inviolable, anything that must be defended with arms) and *ḥarīm* (one's wives and children), probably because the family was regarded as the most sacred thing in the Arab society of the time. When this nuance is taken into account, *ḥtaram* should mean something like: 'to withhold the desire to cause harm because of deep religious or similar feelings'. This nuance has completely disappeared in modern use, however, and it is not clear when exactly that happened. In fact, even its earliest attested uses cannot be claimed with certainty to have had that meaning.

Meaning change does not involve lexis only, but it can also affect sub-lexical components, and for that matter, morphology must also be taken into account when studying language change. It has been argued by many linguists, particularly those who work within the

cognitive linguistic framework, that morphemes as well as the schematic form of derived words usually express meanings, though these meanings are characteristically more abstract than lexical meanings (cf. Bybee 1985, Langacker 1987, 1991, Lakoff 1987, Talmy 2000, among many others). Like lexical meaning, abstract grammatical meaning is also subject to change through extension, shift, re-analysis, or even loss (cf. Fortson 2003, Traugott 2000). The morphological system of the Arabic verb, for instance, can express a variety of grammatical meanings that often determine their argument structure and, thus, have an interface with syntax. Consequently, any change that affects this kind of morphological system can also affect lexical meaning as well as syntactic structure. In this paper, I will illustrate this situation by focusing on the change in the meaning of Form VIII verbs as Classical Arabic (CA) developed into Modern Standard Arabic (MSA).

The paper is constituted of two main sections in addition to this introduction and a conclusion. Section 2 will present the verbal system in different Arabic varieties with the view of setting the scene for later discussions. In this section, we will explain that the focus of this study is not on the development of verbal morphology from CA to the modern colloquial varieties, but rather on the changes that have affected this morphological system during different stages of the standard variety itself, i.e. CA and MSA. Section 3, which constitutes the main contribution of this study, will provide a detailed discussion of the morphological change in Form VIII, which resulted from phonological and/or semantic factors.

2. The verbal system of CA, MSA and the colloquial varieties

Although historical linguists are interested in the study of the mechanisms of language change, they have not always considered seriously the sociolinguistic status of language varieties in their historical development (cf. Romaine 1982). Any diachronic study of the Arabic verbal system, however, cannot afford overlooking the different paths in the development of the Arabic language. This is particularly so because this language is well known for being diglossic and it is not clear when this situation exactly began; some scholars argue that it developed during the early periods of the Muslim empire (e.g. Versteegh 1984: Chap. II) while others claim that it was already characteristic of the pre-Islamic period (e.g. Zwettler 1978: 101). Because of Arabic diglossia, the verbal system of CA followed two paths: one path into the modern dialects, which differ from one variety to another across and beyond the Arab World, and a second path into MSA, a primarily written variety that is spoken natively by no one. Most linguists would show more interest in the first path probably because it is more "natural", but the Doha Historical Dictionary of the Arabic Language is rather concerned with the second. Like other lexicographers, the compilers of this historical dictionary target users of the standard variety in which most learned culture is written. Given that the Arabic learned culture spans over more than a millennium in what is considered to be one and the same language standard, namely *al-Fuṣḥā* (i.e. the pure language), it is no surprise that Arab lexicographers show interest in change in the standard rather than the colloquial.

Arabic, like other Semitic languages, has a basically non-concatenative morphology that weaves roots and patterns into pronounceable words. The tri-consonantal root *K-T-B*, for example, is not a word in itself although it denotes the semantic field of writing; it can be realized as a word only when combined with a (usually) vocalic template, as in *katab* 'to write', *kutib* 'to be written', *kātib* 'writer', *kitāb* 'book', etc. The affixes are quite regular in both form and meaning when combining with other roots to the extent that traditional Arab grammarians refer to them by using the root *F-ʕ-L* 'to do' because of its schematic meaning. Thus, the pattern *faʕal* refers to the perfective form of any tri-consonantal verb, *fuʕil* to the passive of the perfective form of such verbs, etc. In comparison, Western scholars use capital Cs instead to refer to root consonants for practically the same purpose (e.g. McCarthy 1979).

In CA, as in MSA, there are a number of verbal patterns with different schematic meanings. These are called verb forms in the Western tradition of Arabic studies. The most basic one is Form I *CaCaC* for tri-consonantal roots and *CaCCaC* for quadri-literals, both are generally assumed to be non-derived by traditional Arab grammarians, though there are cases that are possibly derived from nouns (e.g. *talfan* 'to give a phone call' from the loan word *tilifūn* 'telephone'). All the other verb forms are augmented by the addition of an affix, a fact which indicates that they are derived. Traditional Arab grammarians generally assume that augmented forms are derived either directly from Form I or indirectly via other augmented forms. For example, Form II CaC_2C_2aC is derived from the basic form by geminating the second root consonant while Form V $taCaC_2C_2aC$ is derived from Form II by prefixing *ta-* to it. If this prefix is attached to the basic form, it would be an inflectional rather than a derivational affix (viz. *taffʕal* 'you do/she does').

The number of augmented forms in Arabic varieties can vary significantly. In CA, for example, 14 augmented forms are usually identified (cf. Wright 1896: I, 29-46) whereas MSA uses only 10 (cf. Ryding 2005: 434). This is clear indication that at least four forms have fallen into disuse as CA developed into MSA. The other forms may not also have the same frequency in the two varieties, nor even the same use for that matter; hence, the significance of scrutinizing their diachronic development. As to the colloquial varieties, the number of verb forms can be extremely reduced. In Moroccan Arabic, for instance, there are only two derived forms: Form II (corresponding to CA Form II) and Form III (corresponding to CA Form V), while the others have completely disappeared because of sound change, or are represented only by a small group of verbs and, consequently, are unlikely to be open to new innovative forms (cf. Harrell 1962: 29-34). In comparison, Gulf Arabic seems to have preserved most of the augmented forms of MSA except form IV, which was lost as a result of the loss of the glottal stop from its phonological system (cf. Qafisheh 1977: Chap. 6). Like other Bedouin dialects, Gulf Arabic seems to be more conservative than those descending from dialects of settled communities (cf. Versteegh 2014: Section 10.3). It is likely, however, that not all the verb patterns are as productive in such Bedouin varieties as they are in MSA or CA.

Regarding the semantics and syntax of verb patterns, it is worth noting that each form is usually associated with a set of senses that determine the argument structure of the clause. Thus, while the basic form verbs can be intransitive, monotransitive or ditransitive, augmented verbs are often associated with a specific clause type given that they are syntactic

and semantic operations on the base form. Form IV, for example, is morphologically derived from Form I by the prefixation of *ʔa-*, and the result of such a derivation is usually the causative form of the Form I verb, as a comparison of *māt* 'to die' and *ʔamāt* 'to kill' shows. The basic form *māt* is intransitive but its augmented equivalent *ʔamāt* is transitive; but when the basic form is transitive, as is the case with *ʕalim* 'to know', the causative is ditransitive; viz. *ʔaʕlam* 'to inform' or *ʕallam* 'to teach'.

Similarly, Form VIII, which is the focus of this study, has its specific syntactic and semantic properties. Form VIII verbs are generally derived from the basic form by the infixation of *-t-* immediately after the first root consonant; e.g. *bāʕ* 'to sell' vs. *btāʕ*[1] 'to buy'. According to traditional Arab grammarians, Form VIII verbs can have up to six different senses (cf. Ibn ʕUṣfūr 1987, II 192-194, Ibn Yaʕīš 2001, IV: 441)[2]. The first meaning can be called "resultative", as exemplified by *rtafaʕ* 'to rise, to climb, to soar' from *rafaʕ* 'to raise, to lift'; it is as if the situation expressed by the derived form of the verb results from that expressed by its basic form. The second meaning that Form VIII verbs can express is the use of some object in some way understood through encyclopedic knowledge, as in *ḥtabas* 'to imprison', that is to take someone as a prisoner. Some Form VIII verbs can also have a reciprocal meaning illustrated by *qtatal* 'to kill each other' derived from *qatal* 'to kill'. Others can refer to the effort made by the subject during the action denoted by the basic form of the verb. For example, the difference between *kasab* 'to earn' and *ktasab* 'to earn' is that the second stresses the role of the agent in the action and, thus, the difference is not truth-functional. As to the remaining two senses, it seems that the tradition cannot distinguish them clearly from the senses of other forms, as is the case of *btasam* and *tabassam* 'to smile', and *xaṭaf* and *xtaṭaf* 'to snatch'. In the first pair, no semantic distinction is detected between the Form VIII and Form V verbs, and the second pair also indicates that the Form I and Form VIII verbs are apparently synonymous. Generally, however, no attempt is made within this tradition to identify the relation between the different senses of a verb form or to explain why they should be expressed by one and the same verb pattern.

By contrast, polysemy in morphology is one of the issues that modern researchers are fascinated with. For these researchers, affixes, just like independent words, usually express different, but related, senses (cf. Copestake & Briscoe 1995, Lehrer 2003, Lieber 2004, Rainer 2014, Rainer et al. 2014, Schulte 2015, among others). Within the cognitive linguistic framework, polysemy is usually explained by the extension of a prototypical sense through metaphor, metonymy or some similar cognitive process. Since verbal affixes express grammatical meanings, they generally operate on the syntactic structure of the clause and, thus, affect the argument structure of the verb. Their syntactic role, however, is rarely uniform precisely because of polysemy, as the discussion of the infix *-t-* of Form VIII in Arabic will illustrate.

[1] This form is pronounced as [ibtāʕ]; the initial i vowel is epenthetic to avoid complex syllable onsets, and will be ignored hereafter.

[2] I would like to thank an anonymous reviewer who drew my attention to al-Tūnī (1997), a study dedicated to Form VIII in the Quran. Though more detailed than what is generally found in Arabic writings on the topic, the study, however, remains faithful to the spirit of the tradition by identifying the same senses mentioned in reference grammars.

The role of the *-t-* infix in Form VIII is primarily to demote the subject and promote the object. A comparison of the examples under (1) illustrates this idea:

(1) a. *ʕazala al-raʔīs-u al-wazīr-a*
'The president dismissed the minister.'
b. *ʕuzila al-wazīr-u*
'The minister was dismissed.'
c. *ʕazala al-wazīr-u nafs-a-hu*
'The minister dismissed himself.'
d. *iʕtazala al-wazīr-u*
'The minister resigned.'

The verb *ʕazal* 'to dismiss' is transitive because it describes an action with an agent and a patient. In (1a), the agent is referred to by the subject NP (marked for nominative case) while the object NP (marked for the accusative) refers to the patient. In (1b), however, the verb is in the passive form, as indicated by the *-u-i-* vocalic melody, and passivization involves the promotion of the NP referring to the patient to the subject position to indicate that the entity is focused. Nonetheless, the agent, though not mentioned, remains in the background as an entity that is distinct from the patient. In comparison, the reflexive pronoun *nafs-a-hu* 'himself' in (1c) indicates that the agent and the patient roles are performed by one and the same entity. Despite that, the basic form of the verb *ʕazal*, by virtue of its semantic content, describes a two-participant event and, consequently, (1c) frames the agent and the patient as conceptually distinct entities. In comparison, the Form VIII *ʕtazal* in (1d) frames the event as a one-participant event, and the action is conceptualized as being performed on the self. In this sense, Form VIII is similar in many respects to the middle voice (cf. Kemmer 1993, Klaiman 1991, Zúñiga & Kittilä 2019).

Like most morphological patterns, Form VIII has uses other than the one discussed in the previous paragraph. Glanville (2018: Chap. 4), one of the rare works on Arabic verbal patterns[3], cites three major uses of this form; these are: subject as beneficiary, actions on the self, and symmetry; which of these is intended will usually be determined by the semantic content of the verb. They are illustrated by the following respective examples:

(2) a. *iqtaṭaʕa ʔarḍ-an*
'He cut out a piece of land for himself.'
b. *irtamā fī al-māʔ*
'He threw himself into the water.'
c. *iǧtamaʕa maʕa ʔaṣḥāb-i-hi*
'He met with his companions.'

In (2a), the verb is derived from the basic *qaṭaʕ* 'to cut', but the result is not an intransitive verb, as is the case in the examples cited in the preceding paragraph. Rather, the derived verb is still transitive and the affix *-t-* adds the meaning that the outcome of cutting will benefit to

[3] An anonymous reviewer thankfully pointed out that Fleisch (1978) and Larcher (2003) also discuss the semantics of Form VIII verbs.

the subject. In comparison, (2b) is a clear case of the middle use already discussed. As to (2c), the verb is derived from *ǧamaʕ* 'to gather' and the resultant meaning is that the subject and the object referents came together. Syntactically, the derived verb is intransitive but the object must be mentioned in the oblique. Glanville (2018: 64-66) also points out that Form VIII verbs can be derived from nouns, given that the form has become established in the grammar as a schema with a set of meanings, though these meanings can be various but related. For example, *ʕtanaq* 'to embrace' comes most probably from *ʕunuq* 'neck' since embracing involves taking and being taken by the neck. The verb is also used to denote adoption of ideas, religions, or theories; but this meaning is apparently an extension from 'to embrace' for reasons that seem obvious. This example is reminiscent of verbs like *ǧtamaʕ* 'to meet' in that the action involves the subject and one or more participants. More will be said later about the various uses of Form VIII and how this polysemy has contributed to semantic change.

In the remainder of this paper, the claim that the meaning and use of Form VIII has undergone some change will be discussed and illustrated on the basis of a list of verbs culled from Baalbaki's (1995) Arabic-English dictionary. This large dictionary was chosen essentially because it is organized on the basis of words rather than roots, as is the usual practice in Arabic lexicography. Since the words follow the alphabetical order, it was much easier to search for verbs with the form *ftaʕal* than would have been the case if a root-based dictionary were used instead. The final list included 529 Form VIII verbs, which were put in a spreadsheet in order to facilitate their arrangement and re-arrangement according to different criteria.

3. Derivational opacity, reanalysis and meaning change

As was mentioned at the outset, this study aims at showing that Arabic words change their meaning in the course of time partly because of opacity in derivational morphology. The relation between morphology and semantics is not a straightforward one and varies according to the theory adopted (see Anderson 2015 and Chapters 19-23 in Part V of Himmisley & Stump 2016). In this paper, however, we will focus more on data than on their theoretical implications and, therefore, we will do our best to couch the description and the argumentation in theory-neutral terms. The issue of opacity/transparency will occupy a central position in the discussion because of its role in semantic change. This section will be divided into three subsections: one on allomorphy, the second on polysemy, and the third on the frequency of the basic form.

3.1. Allomorphy

Allomorphy constitutes a type of morphological change, albeit a small one with non-dramatic consequences. The English simple past marker *-ed*, for example, is usually realized as [d], but can also be realized as [t] in cases like *looked* or as [ɪd] in cases like *wanted*, depending on the preceding consonant. In such a situation, the learner must first make a con-

nection between the three realizations and infer, based on linguistic and contextual clues, that they are allomorphs of the same morpheme, namely the suffix marker of the past tense. However, when the connection between the different allomorphs is no more transparent, this can become a major change leading to the decay or reanalysis of the morpheme. It can result also in the change of the whole paradigm of which that morpheme is an element.

This seems to be the case with the *-t-* of Form VIII. In cases like *manaʕ* 'to prevent' vs. *mtanaʕ* 'to refrain, to abstain', the learner can easily make the connection between the base and the derived forms essentially because the infix surfaces under a form that is identical with the underlying one. When the first root consonant is emphatic, however, the *-t-* gains emphasis by assimilation, as in *ḍarab* 'to hit' vs. *ḍṭarab* 'to be confused'. This kind of allomorphy can be confusing especially that this assimilation excludes the phoneme /r/, which is also emphatic in Arabic, except when followed by the vowel /i/, viz. *ramā* 'to throw' vs. *rtamā* 'to throw oneself'. Similarly, the *-t-* can be realized as [d] when adjacent to /z/, /d/ or /ð/ only, as in *zād* 'to add' vs. *zdād* 'to add to oneself/to be born'. This is apparently a case of voice assimilation although it is restricted to the context of non-emphatic voiced dental sounds. In both emphasis and voice assimilation, the connection between the base and the derived forms can become opaque, thus constituting a challenge for the learner. In yet a third situation, allomorphy can affect not only the affix, but also the stem, making the derivation even more opaque. *ttaxað* 'to assume', for instance, is derived from *ʔaxað* "to take", but the glottal stop assimilates completely to the affix *-t-*. Under this category, we can also include cases of verbs with an initial *t* as a root consonant such as *ittabaʕ* 'to follow', especially that geminates are represented graphically in Arabic by a single letter; viz. اتبع. Similarly, *ddakar* 'to recall' derives from *ðakar* 'to mention', but the affix *-t-* assimilates first to the interdental fricative and the geminate [ðð] is strengthened after that to yield [dd]. (Actually, both *ððakar* and *ddakar* are attested in the language as free or dialectal variants.) Obviously, these morpho-phonological changes obscure the derivational relation between the base and the derived form and, thus, make the learning process much more challenging than would be the case with less opaque derivations. This is true for all learners, but it is more so for L2 learners such as learners of MSA.

The effect of allomorphy on the derivational system will partly depend on its frequency. If only a small set of verbs exhibit differences between the base and the derived forms, allomorphy will probably not have any significant consequences on the morphological category despite its high level of opacity. But when a large number of verbs take a form of the morpheme that is not identical with its underlying representation, the weaker the connection between the different allomorphs is, the more likely the morphological category will be lost. In the case of Form VIII verbs in Arabic, allomorphy seems to have contributed to the opacity of the derivation.

There are 64 cases in which the *-t-* affix occurs under a different form in the list of Form VIII verbs compiled for this study. This figure represents 12.10% of the list; they are divided into five categories exhibited in Table 1.

Table 1. Frequency according to allomorphy

Allomorph/ Grapheme	Frequency	Examples
Geminate /t/	20	*ttabaʕ* 'to follow' *ttaxað* 'to assume' *ttaṣal* 'to be connected to'
Geminate /d/	6	*ddakar* 'to recollect' *ddaxar* 'to save'
Single /d/	11	*zdahar* 'to prosper'
Single /ṭ/	24	*ṣṭafa* 'to choose'
Geminate /ṭ/	3	*ṭṭalaʕ* 'to examine'

The first line shows cases in which the affix *-t-* occurs as part of a geminate, either because the first root consonant is /t/ or because it is a glottal stop or /w/ that assimilates regressively to the affix. In either case, and because the geminate is represented graphically by a single letter, the affix is likely to be opaque to the learner. The same remark holds for the second and the last lines in which the affix surfaces as part of a 'd' or a 'ṭ' geminate, respectively. In the remaining two lines, the affix is represented graphically by a separate letter, but because of progressive assimilation, the letter is different from that found in regular cases, namely ـدـ in the third line and ـط in the fourth. Although 12% does not seem to be a very large percentage, it is large enough to introduce opacity in the derivation and, by consequence, cause confusion to the learner who is initiated to the language typically through the writing mode.

But although allomorphy has a share of responsibility in reducing the transparency of Form VIII derivation, polysemy certainly plays a more crucial role in the semantic opacity. Obviously, when the two factors are combined, the consequences can be drastic, as the discussion below will show.

3.2. Polysemy

Traditionally, polysemy was assumed to be a characteristic of lexical items, and affixes were treated as part of polysemous items. For this reason, the phenomenon of polysemy, and semantic change in general, was rarely discussed in relation to morphology. More recently, however, many researchers have turned to the semantic contribution that affixes bring to the meaning of words (cf. Rainer 2014). In some theoretical frameworks such as Construction Grammar (cf. Booij 2013), for example, an affix is represented as forming a schema together with the grammatical category to which it is attached. The schema has a semantic content, just like any other lexical item in the language. For instance, $[V \_ er]_N$ stands for the combination of a verb and the affix *-er* to form what is called an "agent" noun in English. This schema can denote the agent of some activity, e.g. *writer*, but it could also denote an instrument, e.g. *blender*, or even a theme of an activity, e.g. *bestseller*, etc. This is clear indication that the $[V\_er]_N$ is polysemous.

Similarly, Form VIII is polysemous as a schema. Previously, we pointed out that the function of the *-t-* infix is to turn a transitive verb into an intransitive one by demoting the subject and promoting the object; compare *manaʕ* 'to prevent' and *mtanaʕ* 'to refrain, to abstain' discussed earlier. When a verb is ditransitive, insertion of the infix turns the verb into a monotransitive one, as in the following pair of examples:

(3) a. *bāʕ-a al-tāğir-u al-rağul-a biḍāʕat-an*
'The trader sold the man a merchandise.'
b. *ibtāʕ-a al-rağul-u biḍāʕat-an*
'The man bought a merchandise.'

In (3a), the verb takes a direct and an indirect object, just like its English equivalent. Semantically, the subject is an Agent, the direct object a Theme and the indirect object a Beneficiary. In comparison, the Beneficiary in (3b) is promoted to the subject position while the Agent is backgrounded. While this de-focusing operation is performed lexically in English through the selection of a different verb, it is performed in Arabic morphologically through infixation. In both monotransitive verbs like *manaʕ* and ditransitive verbs like *bāʕ*, the corresponding Form VIII focuses attention on the Patient or the Beneficiary as the initiator of the action. In this sense, the two could be said to express the same meaning, although there is a slight difference between them.

In other cases, however, infixation of *-t-* does not result in any syntactic operation. Such is the case of *btadaʔ* 'to begin' illustrated in these examples:

(4) a. *badaʔ-a al-ʕāmil-u šuγl-a-hu*
'The worker began his work.'
b. *ibtadaʔ-a al-ʕāmil-u šuγl-a-hu*
'The worker began his work.'

In both examples, the verb has the same argument structure, thus indicating that there has been no change in syntactic structure. Semantically, the infix does not seem to bring much to the meaning of the sentence since the activity of beginning work must involve some effort on the part of the agent anyway. Baalbaki (1995) does not provide any equivalent for *btadaʔ* but merely refers the user to the *badaʔ* entry, implying that the two are equivalent. This use of Form VIII is a clear deviation from the one illustrated in the preceding paragraph and, therefore, it must have been a semantic change introduced sometime in the history of the Arabic language.

We have been able to identify eight different uses of Form VIII in the corpus collected, though these should not be treated as tight categories. In addition to the two cases just discussed, which can be considered as the two ends of a continuum, there is a third class of verbs whose Form I and Form VIII are clearly related, though not synonymous. An example that illustrates this class is the pair *raʔā* 'to see, to think' and *rtaʔā* 'to consider'. The fourth category of Form VIII verbs includes verbs for which no corresponding basic form can be identified. As a case in point, *btahal* 'to supplicate' does not seem to be derived from any Form I verb since no such form as *bahal* can be found in the bilingual dictionary from which

the list was culled nor in any other monolingual dictionary of CA. A similar class includes verbs for which there is a corresponding basic verb but with an apparently unrelated meaning. For instance, while both *rāḥ* and *rtāḥ* are attested, the first means 'to leave' and the second 'to rest'. Apparently, the Form VIII verb is derived from the noun *rāḥa* 'rest' rather than from any basic verb form. A sixth class includes Form VIII verbs cited in the bilingual dictionary but for which no entry is cited in monolingual dictionaries of CA or MSA. For instance, Baalbaki (1995) cites *ḥtaðar* with a note of reference to the adjective *ḥaðir* 'cautious', implying that the verb probably means 'to be cautious'. This verb, however, is not cited by dictionaries of CA. The seventh class includes Form VIII verbs that express reciprocity and are, therefore, synonymous with the corresponding Form VI verbs. An example of such verbs is *xtaṣam*; Baalbaki (1995) does not provide any explanation for this verb but merely refers to Form VI *taxāṣam* 'to dispute', implying that the two are synonymous. The last class includes verbs expressing the intensification of an activity. For instance, while the basic form *ḥafā* means 'to welcome', *ḥtafā* means 'to welcome heartily'. As illustrated by the examples, these classes provide clear indication that Form VIII is polysemous.

The eight classes do not all have the same type frequency. By type frequency, we mean the number of verbs in each class, not the frequency of a verb within a given corpus of texts, which is usually called *token* frequency. Table 2 exhibits the frequency of each class.

Table 2. Type frequency of verb classes

Verb class	Example	Frequency	Percentage
Class 1	*mtanaʕ* 'to refrain'	141	26.65
Class 2	*btadaʔ* 'begin'	195	36.86
Class 3	*rtaʔā* 'to suggest'	58	10.96
Class 4	*btahal* 'to supplicate'	51	9.64
Class 5	*rtāḥ* 'to rest'	45	8.50
Class 6	*ḥtaðar* 'to be cautious'	19	3.59
Class 7	*xtaṣam* 'to quarrel'	16	3.02
Class 8	*ḥtafā* 'to welcome heartily'	4	0.75

As can be noticed, the second class ranks first with 195 verbs, covering almost 37% of the list, followed by the first class with 141 verbs. The smallest class includes only 4 verbs while the remaining classes range between 16 and 58. But despite the varying frequencies, the fact that some incompatible classes have more or less the same frequency indicates that the semantics of Form VIII has become very loose. A clear illustration of this point can be provided by a comparison of the first and the second classes. As was explained earlier, while *-t-* affixation in cases like *mtanaʕ* 'to refrain' affects the syntax-semantics of the verb, it does not change much in cases like *btadaʔ* 'to begin'. The question that should be raised in this connection is: how can a language learner connect between the different uses of the infix *-t-*?

In treatments of polysemy, most researchers argue that there is usually a core or "prototypical" sense from which the other senses are derived by metaphor or metonymy or some

other cognitive process (cf. Brugman 1988, Lakoff 1987, Tsohatzidis 1990, Evans 2009, among many others). Regarding the case under study, it seems that Class 1 exemplified by *mtanaʕ* 'to refrain' is the prototype of Form VIII in Arabic. Although this is no place to develop the argument, the fact that many languages have a middle voice through which the subject is defocused and the object focused can be an indication that this is the initial function of this morphological derivational (cf. Kemmer 1990). In other words, all the cases in which no syntactic transformation results from the affixation of *-t-* must have been developed from the prototypical use by extension. From the reflexive use of *-t-* in *mtanaʕ*, for example, the learner may focus on the effort made by the experiencer to refrain from doing something instead of the number of participants in an activity. This attention is transferred later to a verb like *ḥtafā* from Class 8 to intensify the warmth of welcome without any effect on the argument structure of the verb. As long as native speakers are able to infer the right semantics of the non-prototypical uses and their connection to the prototype of an affix, the morphological operation can be claimed to be productive. But when such a connection starts to wane, there is some probability that some change has occurred in morphological derivation as well as in the meaning of the derived forms.

The frequencies in Table 2 above can be interpreted as an indication in this direction. Although Class 1 seems to represent the prototypical use of Class VIII, as argued above, the number of verbs constituting this class is lower than the number of verbs in the second class, which deviates from the prototype. This class constitutes almost 37% of the corpus. Of course, there are several factors that determine the prototypicality of a sense, but frequency is usually regarded as being one of them (cf. Fenk-Oczlon & Fenk 2010). Therefore, the fact that the second class of verbs outnumbers the first could be considered as an indication that Form VIII has changed its prototypical sense or, perhaps more accurately, that it no longer forms a homogenous category from a semantic perspective. This should come as no surprise given that Standard Arabic has been in use as a High diglossic variety for more than a millennium, a time span long enough to account for the reported change. Although there are no studies to my knowledge on the acquisition of the Arabic verbal system by Arab learners (see Zalami 2007 for a review), it seems that even at an advanced level of proficiency, learners are unable to link the various uses of Form VIII, for example, in such a way that these uses form a web of interrelated meanings. They are simply too heterogeneous, both syntactically and semantically, to be included in a single category.

The issue to be considered now is the following: what meaning do language users assign to a Form VIII verb if they are unable to relate it to a basic form through a derivational process? For example, given that there is no attested basic form *bahal*, how do they interpret and store the derived form *btahal* 'to supplicate'? Similarly, how can they link the meaning of *ḍṭarab* 'to be confused' to that of *ḍarab* 'to hit' despite the apparent lack of relation between the two meanings? An attempt to answer these questions will be made after the frequency of the basic form is discussed in the following subsection.

3.3. Frequency of the base form

For a derivation to be productive, both the base and the derived forms must be accessible to the learner and the relation between the two must be transparent to a significant degree. For example, both *manaʕ* 'to prevent' and *mtanaʕ* 'to refrain' are frequent enough in MSA to enable the language learner to establish a certain connection between the two. In comparison, although both *ḍarab* 'to hit' and *ḍṭarab* 'to be confused' are frequent in MSA[4], the fact that the affix has changed into *ṭ* and the meanings of the two verbs are not clearly related has made the two forms unrelated for many speakers. For these reasons, the derivation in the first case will be qualified as transparent while in the second case it is opaque. The more factors there are that intervene to remove the derived form from its base, the more opaque the derivation will be. For this reason, transparency and opacity tend to form a continuum rather than discrete categories.

The intervening factors considered so far are allomorphy and polysemy, and a third factor is the relative frequency of the base. Obviously, if a language user is unable to connect a derived form with its base, it is not clear how that form can be considered derived for him/her. Consider again the example of *btahal* 'to supplicate' from Class 4, for which no corresponding *bahal*[5] can be identified, nor any other base whatsoever. For speakers of MSA at least, there seems to be no sense in claiming that this verb is constituted of a base and an affix, and that its meaning is the result of combining the meanings of its constituents. Therefore, it is very likely that cases like this are learned as non-derived verbs and that the *t* is reanalyzed as a root consonant rather than an affix. Class 5 may also be behaving in the same way. As explained above, this class includes verbs whose Form I and Form VIII do not share the same meaning, at least in MSA usage, though they share the same root consonants. For instance, *btaɣā* is derived from *baɣā*, both of which mean 'to seek' in CA; but in MSA, *baɣā* is usually used with the meaning of 'to assault'[6]. Therefore, speakers of MSA are unlikely to consider the second as the base of derivation for the first and, consequently, the *t* of *btaɣā* is perhaps not treated as an affix. Given that the two classes of verbs include 96 cases and form more than 18% of the corpus (see Table 2 above), we can easily imagine the impact such cases must have had on the internal consistency of Form VIII as a morphological category.

Frequency does not involve only these two classes but cuts through all the others as well. Even Class 1 includes cases in which the corresponding Form I verb is of low frequency in MSA and may not be familiar to a large number of speakers. For instance, *ntaʕaš* 'to become refreshed' should correspond to *naʕaš*, a form that does exist in CA with the meaning of 'to

[4] *Ḍarab* occurs around 32 thousand times in ArabiCorpus while *ḍṭarab* occurs about one thousand times.

[5] While *btahal* occurs 145 times in ArabiCorpus, a search for *bahal* does return some 9 instances. A close scrutiny, however, indicates that these are colloquial forms from Levantine Arabic that are combinations of the preposition *b-* 'with' and the demonstrative *hal* 'this'/'these'.

[6] This is particularly the case in parts of the Arab World where *bɣa* is used in the colloquial variety with the meaning of 'to want' such as North Africa. Apparently, MSA users avoid such forms in order not to be suspected of wrong use of the language. Where this is not the case, however, as in the Middle East, *baɣā* is used with same meaning, as an inspection of the 907 instances in ArabiCorpus shows.

raise' but which is almost inexistent in MSA; the ArabiCorpus, for example, does not include any instance of this form. Baalbaki (1995) does cite *naʕaš* but merely refers to Form IV *ʔanʕaš* for explanation, implying that the two have more or less the same meaning. For MSA speakers, *ntaʕaš* is more linked to *ʔanʕaš* than to any other form, but it is not clear how one could be derived from the other. In other cases, the base form may be familiar to MSA speakers but with a different meaning. As a case in point, both *nahā* 'to prohibit' and *ntahā* 'to finish' are quite frequent in modern usage, but the first is linked to the action noun *nahy* 'prohibition' and the second to *nihāya* 'end, termination' or *ntihāʔ* 'completion'. In CA, however, all these seem to be connected; specifically, *ntahā* used to denote restraining oneself from doing or enjoying something, and the meaning of putting an end to an activity is a mere extension of this sense by implication. In comparison, the two senses are separate in MSA. Therefore, we must conclude that the organization of the lexicon of MSA must be different from that of CA, a conclusion that is not surprising given that the second was spoken by native speakers while the first is used by second language learners only (cf. Khamis-Dakwar & Froud 2019, Froud & Khamis-Dakwar 2021). Some words are often represented in the mental lexicon of second language learners as simplex even when they are complex (cf. Milton 2009: 103-105). Although it was not possible to measure the frequency of the base form of all the verbs in the corpus due to ambiguities in the classification, a significant number of cases seem to lack a transparent connection between the base and the derived forms.

Obviously, the more opaque the relation is between the base and the derived forms, the more likely the derivational process will be obscured and, ultimately, lost. Semantically, once the two forms are no more morphologically linked, each will develop its own meaning separately from the other. To take the examples of *btahal* 'to supplicate' and *ntahā* 'to finish' once again, each of them has developed a separate meaning. In particular, the first stands now on its own since its base was lost in the course of language development and, consequently, can be argued to be synchronically underived. As to the second, it stands somewhere between transparent verbs and completely opaque verbs. On the one hand, it seems to be unrelated to *nahā* 'to forbid' for many speakers of MSA, but on the other, it is still connected to the noun *nihāya* 'end', which does not include the affix *-t-*. The fact that *ntahā* and *nihāya* are semantically related is likely to encourage the analysis of the first as a Form VIII verb, possibly derived from the noun form instead of the putative base *nahā*. (It should be recalled that many Form VIII verbs are derived from nouns, not necessarily Form I verbs, as explained in Section 2 above.) If this is indeed the case, the meaning of *ntahā* in MSA should not include any reference to self-restraint, which apparently it used to have in CA by virtue of its derivation from *nahā*. This remark holds for most cases exhibiting some degree of opacity due to allomorphy, polysemy or infrequency of the basic form.

One final caveat, however, is in order. Speakers of Arabic today are of varying degrees of proficiency and may have different experiences with written Arabic. They may even hold different perceptions and attitudes toward the various styles and usages, which could manifest in their prescriptive views. Therefore, unless deep investigation has been carried out into the mental lexicon of different categories of speakers, our understanding of the degree of semantic change in the verbal morphology of the language will remain imprecise. Some

change has occurred, nonetheless, though it needs to be studied in language use rather than on the basis of individual intuitions only.

4. Conclusion

It has been argued in this paper that Form VIII verbs show various degrees of transparency/opacity depending on at least three factors: allomorphy, polysemy, and frequency of the base. If transparency/opacity can be measured against the intuitions of modern day speakers of the language, no access is possible to past speakers. One consequence of this fact is that we may never know whether and to what extent a morphological derivation was productive in the past and, if some change happened in this regard, when it happened exactly. To take the example of *ḥtaram* 'to respect' again as cited in the introduction, we will probably never know whether the first recorded occurrence of this verb was treated as a form derived from *ḥarum* 'to be forbidden' or not, and if yes, when it stopped to be so as is the case for modern speakers. This being said, the study of texts from a given period could turn out to be helpful in determining the frequency of the derived forms and their putative bases. As argued in this paper, when a base is frequent with a given sense, it is more likely that part of that sense will be maintained in the derived form than when it is not; compare *manaʕ* vs. *mtanaʕ* and *mahan* 'to serve' vs. *mtahan* 'to practice a profession'. Unlike *manaʕ*, *mahan* is very archaic and is, therefore, unknown to most users MSA. Thus, those who use *mtahan* today are unlikely to think of it as derived from *mahan* and, consequently, would not include 'service' as a component of its meaning much like they would exclude 'forbidding' from the meaning of *ḥtaram*.

References

Anderson, S. 2015. Morphological change. In Bowern, C. & Evans, B. (eds.), *The Routledge handbook of historical linguistics*, 264-285. London – New York: Routledge.

Baalbaki, R. 1995. *Al-Mawrid: A modern Arabic-English dictionary*. Beirut: Dār al-ʕIlm li-l-Malāyīn.

Bahloul, M. 2008. *Structure and function of the Arabic verb*. London – New York: Routledge.

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

Booij, G. E. 2013. Morphology in Construction Grammar. In Hoffmann, T. & Trousdale, G. (eds.), *The Oxford handbook of Construction Grammar*, 255-273. Oxford: Oxford University Press.

Brugman, C. 1988. *The story of 'over': Polysemy, semantics and the structure of the lexicon*. New York: Garland Press.

Bybee, J. L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam – Philadelphia: Benjamins.

Copestake, A. & Briscoe, T. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12. 15-67.

Evans, V. 2009. *How words mean: Lexical concepts, cognitive models and meaning construction*. Oxford: Oxford University Press.

Fenk-Oczlon, G. & Fenk, A. 2010. Frequency effects on the emergence of polysemy and homophony. *International Journal "Information Technologies and Knowledge*" 4(2). 103-109.

Fleisch, F. 1979. *Traité de philologie arabe*. Vol. II: *Pronoms, morphologie verbale, particules*. Beyrouth: Dar El-Machreq.

Fortson, B. W. 2003. An approach to semantic change. In Joseph, D. & Janda, R. (eds.), *The handbook of historical linguistics*, 648-666. Malden etc.: Blackwell.

Froud, K. & Khamis-Dakwar, R. 2021. The study of Arabic language acquisition: A critical review. In Ryding, K. C. & Wilmsen, D. (eds.), *The Cambridge handbook of Arabic linguistics*, 48-82. Cambridge: Cambridge University Press.

Glanville, P. J. 2018. *The lexical semantics of the Arabic verb.* Oxford: Oxford University Press.

Harrell, R. 1962. *A short reference grammar of Moroccan Arabic*. Washington: Georgetown University Press.

Himmisley, A. & Stump, G. 2016. *The Cambridge handbook of morphology*. Cambridge: Cambridge University Press.

Ibn ʕUṣfūr, A. 1987. *Al-Mumtiʕ fī al-taṣrīf*. Ed. by F. Qabāwa. Bayrūt: Dar al-Maʕrifa.

Ibn Yaʕīš, M. 2001. *Šarḥ al-Mufaṣṣal*, vol. 4. Dār al-Kutub al-ʕIlmiyya.

Kemmer, S. 1993. *The middle voice*. Amsterdam – Philadelphia: John Benjamins Publishing Company.

Khamis-Dakwar, R. & Froud, K. 2019. Diglossia and language development. In Al-Wer, E. & Horesh, U. (eds.), *The Routledge handbook of Arabic sociolinguistics*, 300-313. Abingdon – New York: Routledge.

Klaiman, M. H. 1991. *Grammatical voice*. Cambridge: Cambridge University Press.

Kotowski, S. & Plag, I. (eds.). 2023. *The semantics of derivational morphology: Theory, methods, evidence*. Berlin – Boston: De Gruyter.

Lakoff, G. 1987. *Women, fire and dangerous things: What categories can reveal about the mind*. Chicago: Chicago University Press.

Langacker, R. 1987. *Foundations of cognitive grammar I: Theoretical pre-requisites*. Stanford: Stanford University Press.

Langacker, R. 1991. *Foundations of cognitive grammar II: Descriptive application*. Stanford: Stanford University Press.

Larcher, P. 2003. *Le système verbal de l'arabe classique*. Aix-en-Provence: Publications de l'Université de Provence.

Lehrer, A. 2003. Polysemy in derivational affixes. In Nerlich, B. & Todd, Z. & Herman, V. & Clarke, D. D. (eds.), *Polysemy: Flexible patterns of meaning in mind and language*, 217-232. Berlin – New York: Mouton de Gruyter.

Lieber, R. 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press.

McCarthy, J. 1979. *Formal problems in Semitic phonology and morphology*. Cambridge, Massachusetts: MIT. (Doctoral dissertation.)

Milton, J. 2009. *Measuring second language vocabulary acquisition.* Bristol – Buffalo – Toronto: Multilingual Matters.

Muʕğam = Muʕğam al-Dawḥa al-Tārīxī li-l-Luɣa al-ʕArabiyya. (https://www.dohadictionary.org/) (Accessed 2023-05-10.)

Qafisheh, H. 1977. *A short reference grammar of Gulf Arabic*. Tucson: University of Arizona Press.

Rainer, F. 2014. Polysemy in derivation. In Lieber, R & Štekauer, P. (eds.), *The Oxford handbook of derivational morphology*, 338-353. Oxford: Oxford University Press.

Rainer, F. & Dressler, W. U. & Gardani, F. & Luschützky, H. C. 2014. Morphology and meaning: An overview. In Rainer, F. & Gardani, F. & Luschützky, H. C. & Dressler, W. U. (eds.), *Morphology and meaning*, 3-46. Amsterdam: Benjamins.

Romaine, S. 1982. *Socio-historical linguistics: Its status and methodology*. Cambridge University Press.

Ryding, K. 2005. *A reference grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

Saad, G. N. 1982. *Transitivity, causation and passivization: A semantic-syntactic study of the verb in Classical Arabic*. London – New York: Routledge.

Schulte, M. 2015. *The semantics of derivational morphology: A synchronic and diachronic investigation of the suffixes* -age *and* -ery *in English*. Tübingen: Narr.

Talmy, L. 2000. *Toward a cognitive semantics*. 2 vols. Cambridge – London: The MIT Press.

Traugott, E. C. 2000. Semantic change: An overview. In Chen, L. & Sybesma, R. (eds.), *The First Glot International State-of-the-Article Book: The latest in linguistics*, 385-406. Berlin: Mouton de Gruyter.

Tsohatzidis, S. L. (ed.). 1990. *Meanings and prototypes: Studies in linguistic categorization*. London – New York: Routledge.

al-Tūnī, M. Z. Ḥ. (1997). Ṣīγat *iftaʕal* fī al-Qurʔān al-Karīm: Dirāsa luγawiyya. *Mağallat Kulliyyat al-Tarbiya (al-Qism al-ʔAdabī)* 3(3). 1-120.

Versteegh, K. 1984. *Pidginization and creolization: The case of Arabic*. Amsterdam – Philadelphia: John Benjamins.
Versteegh, K. 2014. *The Arabic language*. Edinburgh: Edinburgh University Press.
Wright, W. 1896. *A grammar of the Arabic language*, vol. 1. Cambridge: Cambridge University Press.
Zalami, F. B. 2007. First language acquisition. In Versteegh, K. & Eid, M. & Elgibali, A. & Woidich, M. & Zaborski, A. (eds.), *Encyclopedia of Arabic language and linguistics*, vol. 2, 96-104. Leiden – Boston: Brill.
Zúñiga, F. & Kittilä, S. 2019. *Grammatical voice*. Cambridge: Cambridge University Press.
Zwettler, M. 1978. *The oral tradition of Classical Arabic poetry: Its character and implications*. Columbus: Ohio State University Press.

DOI: 10.14746/linpo.2025.67.2.3

Synchronic morphological parsing and suffixal productivity in the Hittite adjective: Considerations on some illustrative examples*

Marianna Pozza

Sapienza University of Rome
marianna.pozza@uniroma1.it | ORCID: 0000-0002-2105-3262

Abstract: The purpose of this article is to analyse some Hittite adjectives, to discuss the different options concerning their morphological parsing – and, if available, their etymology – and to clarify the issues behind the productivity of some specific suffixes.

Keywords: Hittite, derivational morphology, morphological productivity, adjective, Proto-Indo-European, etymology

0. Premise

The purpose of this contribution is to analyse, from the perspective of morphological parsability – disregarding any diachronic consideration – some Hittite adjective formations of problematic interpretation, for which it does not seem possible to perform a parsing that clearly isolates the derivational morphemes and precisely defines the class membership of the base from which they derive.[1] The analysis presented here takes into account the functionalist theoretical framework of Natural Morphology (cf. Mayerthaler 1980, Dressler et al.

* The present research has been carried out within the PRIN PNRR Project 2022 "The beginning of the Armenian, Syriac (and Arabic) grammatical traditions and the classical Greek grammar" coordinated by Giancarlo Schirru (P2022LWSYY). I would like to thank the two anonymous reviewers for their thoughtful comments and efforts towards improving this article. Any remaining inaccuracies are my own responsibility.

[1] For the controversy between the morpheme-based ("parsability") and the word-based ("non-parsability") approach see, e.g., Hill (2020: e52), Audring & Masini (2019: 15-16), Haspelmath & Sims (2010: 40-53), Bybee (1995: 426, fn. 1; 1985: 127-129), etc. However, according to recent psycholinguistic studies (cf. Hill 2020: e53 for the bibliographical references), speakers would be capable both to process morphologically complex forms as a whole and to parse inflectional forms into smaller units (cf., among others, Burani & Thornton 2003).

LP LXVII (2), 2025. © The Author(s). This is an open access article licensed under the CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en/).

1987, Kilani-Schoch 1988, Bybee 1995, Dressler 2000, etc.), which, within the *continuum* between derivation and lexicon – whose boundaries are notoriously blurred – focuses on words that, although characterised by a semi-transparent internal structure and analysable through diachronic processes, can synchronically be considered as stored in the speaker's lexicon when the diagrammatic relationship between semantic and morphological motivation has been lost. As is well known, current debates in linguistic theory highlight a fundamental opposition regarding how speakers process inflectional forms. One school of thought, characterised by a morpheme-based approach, suggests that an individual's mental dictionary largely comprises morphemic elements and the combinatory principles that enable them to construct inflected forms (cf., e.g., Bruening 2018). In contrast, another perspective (cf., e.g., Blevins 2016) – commonly described as word-based – holds that speakers store entire words in their memory and rely on these word sets as templates whenever a particular form is not directly accessible. However, the present study aims to address several problematic cases concerning morphological parsability in Hittite adjectives and to discuss various interpretative options without necessarily taking a stance on either theoretical perspective. As will be shown below, this neutrality does not preclude the possibility of situating certain words at the boundaries, or elsewhere, within the *continuum* between derivation and lexicon. In the present article, the term primary (adjective) reflects a broader problem: it is used non only referring to the general meaning of 'simple', 'underived' (i.e., mono-morphemic, apart from its ending), but also – as we will see later – to words exhibiting the structure *root* + *suffix* + *ending* (where the root is not attested). In both cases, reference is made to adjectives that are stored in the lexicon as 'primary words' because they cannot be formed via synchronically productive rules (albeit showing some traces of internal structure).

After a brief overview of adjective formation in Hittite (§ 1) and a discussion about synchronic morphological parsing of complex words (§ 2), the article will examine several noteworthy specific formations. The following paragraphs will therefore focus, respectively, on the problematic suffix *-ena-/-ina-* (§ 2.1), in order to determine whether it attaches to a coherent class of nouns and in which cases it is actually segmentable within a word; on some underived *-i-* and *-u-*stem adjectives (§ 2.2); and, finally, on certain formations in *-ant-* (§ 2.3) and in *-want-* (§ 2.4), both well attested in Hittite.

1. Adjective formation in Hittite

As is well known,[2] Hittite adjectives can be underived, in which case they mostly end in *-a* (cf. *arawa-* 'free', *kappi-* 'small, little', *nakkī-* 'heavy, difficult', *nēwa-* 'new', *tepu-* 'small, little', etc.), or derived, formed through the addition of one or more derivational suffixes. Loanwords are usually included in the category of underived adjectives,

[2] EHS: 160ff., Berman (1972), Hoffner & Melchert (2008: 51ff.), Francia & Pisaniello (2019: 26). The most comprehensive study of how Hittite nominal stems are formed is found in EHS. Additional insightful analyses of specific stem categories can be found in the works of Weitenberg (1984) and Rieken (1999).

because they are not characterised by Hittite derivational suffixes (Hoffner & Melchert 2024: 86).

Among the main suffixes that form Hittite adjectives, the following can be mentioned (cf. Hoffner & Melchert 2008: 54 ff.): *-ala-*, which forms adjectives from nouns (cf. *liššiyala-* 'pertaining to liver' < *lišši-* 'liver' etc.) or from adverbs (cfr. *tuwala-* 'far' < *tuwa* 'at distant' etc.); *-alla-* and *-alli-*[3], which form denominal adjectives (cfr. *annalla-* 'maternal' < *anna-* 'mother', *attalla-* 'paternal' < *atta-* 'father' etc.); *-aššа/i-*, which forms denominal adjectives denoting appurtenance (cf. ^{URU}*Tarḫuntašša-* '(city) of *Tarḫunta*' etc.); *-iya-*, which forms denominal and deadverbial adjectives (cf. *išpantiya-* 'nocturnal' < *išpant-* 'night' etc.); *-ili-*, which forms adjectives from different bases (cf. *karuili-* 'previous, past' < *karū* 'previously, in the past', *tarḫuili-* 'strong, powerful' < *tarḫu-* 'prevail, conquer' etc.); *-want-*, which forms possessive adjectives from nouns (cf. *ešḫarwant-* 'bloodstained' < *ešḫar-* 'blood', *pittuliyawant-* 'anxious' < *pittuliya-* 'fear, tension' etc.) and from verbs (*armaḫḫuwant-* 'pregnant' < *armaḫḫ-* 'make pregnant' etc.)[4]; and, finally, *-zzi*(*ya*)- which forms adjectives from locative adverbs (cf. *appezzi*(*ya*)- 'last, most recent' < *āppa* 'behind, after,' etc.). The situation concerning the suffix -(*a*)*nt-* is complex; from a purely synchronic perspective, it can be considered as a multifunctional suffix.[5] It forms denominal possessive adjectives (such as *perunant-* 'rocky' < *peruna-* 'rock,' *irmanant-* 'ill, suffering from a disease' < *irman* 'disease,' etc.) and has an individualising function (cf. *ḫamešḫant-* 'the (next) spring' < *ḫamešḫa-* 'spring,' etc.), as well as an "empty" function, because forms characterised by this suffix show no semantic difference compared to those without it (cf. *ikuna-* and *ikunant-* 'cold,' *gaena-* and *gaenant-* 'relative,' *ḫappina-* and *ḫappinant-* 'rich,' *dannara-* and *dannarant-* 'empty, smooth,' etc.). [6]

Other strategies for adjective formation (although adjectival attestations are less frequent than nominal or verbal ones) include reduplication (cf. *walliwalli*(*ya*)- 'fast (?), strong (?)' < *walli-* 'glory, pride', [7] etc.) and compounding (cf. *dāyuga-* 'of two years,' constructed on **dā-* < **dwoyo-* 'two' – cf. *dān* 'for the second time' – and *yuga-* 'year/season,' cf. Hoffner & Melchert, 2008: 153; EHS: 116ff.).

[3] Cf. Melchert (2005: 455-456) for the discussion on "*i*-mutation".

[4] On *-want-*, cf. Oettinger (1988, 2022), Frotscher (2013, 2017), Maier (2013), and Rieken & Sasseville (2014).

[5] Cf. Hoffner & Melchert (2008: 55 fn.10). See also Hoffner & Melchert (2008: 55, note 10). For more detail, refer to Frotscher's doctoral thesis (2013), which, in addition to exploring various properties (as will be discussed further), distinguishes denominal adjectives (e.g., *laḫlaḫḫimant-* 'excited' < *laḫlaḫ(ḫ)ima-* 'excitement'), denominal nouns (e.g., *ḫuḫḫant-* 'grandfather' < *ḫuḫḫa-* id.), and deadjectival formations (e.g., *antarant-* 'blue' < *antara-* id.).

[6] In these specific cases, the -(*a*)*nt-* suffix is traditionally considered semantically empty, because forms bearing it exhibit no difference in meaning relative to those without it. Regarding the (poly)functionality of -(*a*)*nt-*, see Dardano (2010), Frotscher (2013), Melchert (2017a), Rieken (2017), and Goedegebuure (2018).

[7] Hittite *walliwalli*(*ya*)- 'impetuous, stormy, strong' is probably related to *walli-* 'glory, pride,' although the formation is not entirely clear: the noun only appears in the genitive singular *walliyaš pedan* 'place of glory' (HEG W-Z: 260).

2. Synchronic morphological parsing of complex words and etymology

There are numerous discussions regarding the evaluation of the typological category of the adjective[8] in the Indo-European stage, as a lexical class separate from that of the noun, from which the former supposedly differs by virtue of a richer and more productive gender inflection (a contextual category for adjectives and an inherent one for nouns, where the former is mostly a derivational category relative to the latter). Participles themselves do not constitute a separate class but are often understood as deverbal adjectives. Moreover, a closely related aspect to the evaluation of adjectival typology – but which will not be addressed explicitly in this contribution – is the debate on the Caland system and its suffixes (*-*u*-, *-*ro*-, *-*mo*-, *-*nt*- and, perhaps, *-*i*-).[9] In the case of Hittite, a language in which its functioning is relatively straightforward, reference may be made to Hoffner & Melchert (2024: 85ff.), who note the difference between the suffixation process by addition (from *išpant*- 'libate, pour' → *išpant-uzzi*- 'libation' → *išpant-uzzi-aššar* 'libation vessel', etc.) and by substitution, which is mainly found when the base is an adjective (*park-u*- 'high', *parg-ašti*- 'height', *park-nu*- 'to elevate', *park-ešš*- 'to become high', etc.) and which reveals traces of the more general Indo-European Caland system.

From a synchronic perspective, however, suffixation by addition was likely the regular model for Hittite speakers. In contrast, suffixation by substitution was perceived as synchronically irregular (Hoffner & Melchert 2024: 85). This would explain suffixed formations such as *ḫatku*- 'narrow' > *ḫatku-ešš*- 'to become narrow,' *uktūri*- 'durable' > *uktūri*(*y*)-*aḫḫ*- 'to make durable,' *nakkī*- 'important' > *nakki*(*y*)-*atar* 'importance, dignity,' *mayant*- 'young' > *maya*(*n*)*d-atar* 'youth' and *mayant-aḫḫ*- 'to make young.' More generally, the typological aspect of the debate on the root has interesting consequences for the reconstruction of Indo-European. The two derivational systems are not mutually exclusive (cf. Alfieri 2023: 262), insofar as not all derived formations are better explained within one scenario rather than the other, and the debate among scholars remains open. Likewise, discussions on the origin of the Indo-European adjective tend to proceed in two opposite directions: some authors argue that quality values were encoded as nouns in the lexicon (and that consequently, in the Indo-European stage, nouns and adjectives belonged to a single lexical class; see, among others, Balles 2006 and 2008); others have suggested that Indo-European adjectives manifested

[8] Cf., among others, Comrie (1997: 101ff.), Szemerényi (1985: 191ff.), Meier-Brügger (2002: 292ff.). For a recent account of the adjective as a word class see Beck (2023), and for an overview of the adjective from a typological perpective cf. Dixon (2004).

[9] The literature on this issue is vast. For the main discussions see Risch (1974: 65-112), Meissner (1998, 2006), Meier-Brügger (2002: 292ff.), Stüber (2002), Rau (2009: 67-75), Dell'Oro (2015), Oettinger (2017), Alfieri (2023) and see Dardano (2007) for an analysis of the Hittite material. For a summary of the issue, especially regarding the morphemes that, over time, have been added to the list of the more traditional Caland suffixes, see Bichlmeier (2015: 258), according to whom the suffix *-*e/ont*- "is probably not identical to the suffix of the active participles in PIE *-*e/ont-/-n̥t*-". The question of the identity of these suffixes is still being disputed (cf. Lowe 2014). For the controversial state of *-*i*-adjectives, cf. the bibliography quoted in in Alfieri & Pozza (2024), in particular Nussbaum (1976, 2014), Tronci (2000), Bozzone (2016), Grestenberger (2013, 2014, 2017), Lundquist & Yates (2018: 2115), and Höfler (2022).

a verbal orientation (they would be deverbal formations, participles, etc.) rather than a nominal one (e.g., Alfieri 2009 and Bozzone 2016).

Despite some previous works having addressed the relationship between the Caland system and Hittite (notably Dardano 2007, see also § 2.3), the morphological decomposition method used to discuss the cases presented in this study and the proposed analysis of certain forms as "primary" or "derived," follows the principle[10] according to which morphologically segmentable complex words are those formed on attested bases through word-formation rules that are fully transparent both semantically and formally. Derivatives that do not meet these criteria – although possessing some internal structure – are instead considered as stored in the lexicon (consider adjectives representing fossilisations of ancient participles, such as Italian *lucente* 'bright' < Latin *lucens* < *lucēre* 'brighten', which cannot be synchronically derived from an existing verbal base, even though the formation of active participles in -V*nte* is a productive rule in Italian, cf. Bozzone 2016).[11] Therefore, semi-transparent formations that, while analysable diachronically, cannot be segmented by applying derivational rules, will be regarded as lexicalised elements (cf. Bauer 2001: 27, 43)[12] and as underived, thus "primary" formations ("simple words are the hard core of storage," cf. Mayerthaler 1987: 46) not subject to further morphological parsing by the speaker. The frequency with which a given form is attested is also, as is well known, an additional factor that can potentially transform originally derived words into independent lexical items (cf. Bybee 1985: 133; 1995: 429). However, it should be recalled that when dealing with a *corpus*-language such as Hittite, it is virtually impossible to assess token frequency on the basis of the available textual record.

Usually (based on a more general systematisation criterion), derivatives characterised by the same suffix are grouped in reference grammars (cf. Hoffner & Melchert 2008 and 2024), regardless of whether they result from productive rules. Nonetheless (cf. Alfieri & Pozza 2024: 154), in line with contemporary morphological research, considering adjectives to be lexicalised when rules cannot synchronically process their structure avoids the necessity to postulate unattested words (a reconstruction which is, however, entirely relevant from a diachronic viewpoint) and to deduce the word-class status based solely on the suffix type. In Hittite, most suffixes are not exclusively affixed to a single type of base (cf. Hoffner & Melchert, 2008: 54 ff., 2024: 88 ff.).

[10] Cf. Mayerthaler (1980), Dressler *et al.* (1987), Bybee (1995), in addition to the references quoted in footnote 12.

[11] See also examples such as Italian *fantasista* < *fantasia*, *animalista* < *animale*, or English *murderer* < *murder*, *worker* < *work* compared to *ametista, pista, spider, hammer*, etc. A similar method of composition relies on the notion of productivity (Bauer 2001, 2005), which can only be indirectly assessed in ancient languages. On affix productivity in closed corpus languages, see also Panagl (1982). For productivity and diachrony, see Sandell (2015).

[12] For more on lexicalisation processes of adjectives in different ancient Indo-European languages and general theoretical aspects of part-of-speech analysis from a typological and comparative perspective, see numerous works by L. Alfieri (especially Alfieri 2014, 2016, 2021), recently cited in Alfieri & Pozza (2024). On factors triggering lexical storage of derived words cf. Dressler et al. (1987), Bertram et al. (2000), Aronoff & Anshen (2001), Bell & Schäfer (2016). See also Lipka (1994).

For example, the suffix -(*a*)*nt*- attaches to verbal bases (*akkant*- 'dead' < *akk*- 'to die'), to nouns (*irmanant*- 'ill, suffering from a disease' < *irman*- 'disease'), and to adjectives (*ikuna*- and *ikunant*- 'cold') – indeed, it is not always easy to establish a clear boundary between denominal formations in -(*a*)*nt*- and participles – and the suffix -*want*- attaches both to nominal bases (*ešḫar-want*- 'bloodstained' < *ešḫar*- 'blood') and to verbs (*kartimmiya-want*- 'angry' < *kartimmiya*- 'to be angry').[13]

Moreover, Hittite is a language that had profound contacts (starting with its mixed graphic system, both logographic and phonographic) with other languages of the Ancient Near East, not only those of the Indo-European family.[14] This characteristic complicates the analysis of derivational affixes, whose nature and existence are not always easy to establish, as will be discussed in § 2.1). Furthermore, it cannot be excluded that unattested forms in Hittite might have been present in lost or still undiscovered texts, just as related forms may be documented in other Anatolian languages. However, even if a base form is documented in another Anatolian language, this alone is insufficient to hypothesise its (potential) presence in Hittite. When considering synchronic productivity, it is essential to keep in mind that the reconstruction of a base form that generated a (attested) derived form can be performed from a diachronic perspective – reconstructing a proto-form not necessarily attested – but this does not always synchronically clarify the word-formation processes. Therefore, as will be explicitly observed later, it may lead to assessing the lexeme as an indivisible element. Finally, it should be emphasized that the attested Hittite corpus may not fully capture the polysemy of either base or derived lexemes. In other words, without access to the complete semantic range of the items in question, it is impossible to determine whether a given derivational relationship was genuinely affected by semantic drift.

Essentially, from a historical perspective (see §§ 2.3 and 2.4), it is legitimate to analyse adjectives such as **idālu*- 'bad, evil' and *le/iliwant*- 'fast, winged, urgent' respectively as derivatives in -*u*- and in -*want*- (or in -*ant*-) from bases such as **idāl*- (< **edwal*-, cf. Cun.Luw. *ādduwāl*(*i*)- id.) and **le/ili*- (or **le/iliw*(*a*)-) (cf. EDHIL: 421, 525; HED E/I: 493, L: 84-85; HEG A-K: 443, L-M: 58). However, the bases marked with an asterisk are reconstructed and not (yet) attested, which leads to them being considered as non-existent, and consequently the formations in some way connected to such bases being classified as primary, despite their potential analysability diachronically.

Consider, for example, *iyatnuwant*- 'luxuriant', probably derived[15] from the oblique stem of a fossilised verbal noun, *iyatar*, indicating 'growth', 'fertility' (especially vegetal) + the suffix -*want*- (or participle from **iyatnu*-, cf. EHS: 568, albeit with doubts) and

[13] Furthermore, within nominal morphology, consider the suffixes -*att*-, -*eššar*-, and -*ima*-, which attach both to verbs and adjectives, or the suffix -*ātar*, which can attach to nouns, verbs, and adjectives (for -*ātar* and -*eššar* see, recently, Pozza & Fagiolo 2023 with the bibliography therein). For a precise classification of clearly denominal or deverbal forms (in -*ant*-/-*want*-), see Frotscher (2013: 39 ff.).

[14] See, among others, Alfieri & Pozza (2024) and the bibliography therein, especially the numerous works by P. Dardano on this topic, beginning with Dardano (2011).

[15] Cf. Watkins (1979: 282), who interprets *iyatar* as an abstract denoting 'movable wealth', literally 'that which goes, that which moves' (contra, HED E/I: 348-349; 352). For a different interpretation, see the discussion in Rieken (1999: 254ff.).

iyatniyant- 'cultivated', 'growing' (an adjective in *-nt-* from a base with an extension in *-na-/-niya-* or a participle of an unattested denominal **iyatniya-*, cf. EHS: 101; 568, HEG A-K: 348, HW^2 I: 32), showing clear correlation with *iya-* 'to walk', in the specific sense of 'to grow'. However, as noted by Puhvel (HED E/I: 352), the synchronic relation with ("the living paradigm of") *iya-* does not occur, due to the lack of assimilation of *-tn-* into *-nn-* as expected in the oblique cases of verbal nouns in *-atar* (genitive **iyatnaš*, instead of the expected ***iyannaš*).[16] In addition to the problem of lack of assimilation, *iyatar* is frequently attested without the final vibrant, as observed by Starke (1990: 473) and Rieken (1999: 255), which makes the connection with *iya-* less obvious. Furthermore, the attestation (old Hittite in middle script) of the singular genitive *iyataš* would support the hypothesis (EDHIL: 380) that *iyata-* represented the base for *iyatar* and not simply that the former was a variant lacking /r/.[17]

The relationship between *iyatar* and *iya-*, essentially, can be etymologically founded. From a synchronic perspective, the association between the two forms based on the productivity of the suffix of verbal abstracts is impossible, leading to the consideration of *iyatar* as a primary noun endowed with a semi-transparent internal structure.

2.1. The case of the suffix *-ena-/-ina-*

Some lexical items may lend themselves to more than one interpretation depending on the criterion adopted for morphological segmentation: consider, for instance, the case of the obscure suffix *-ena-/-ina-*,[18] whose presence is not always easily identified in some Hittite words (as discussed in Pozza 2023). Additionally, a considerable number of forms ending in *-ena-/-ina-* (cf. Jie 1994: 14-15) lack etymology, partly because their meaning cannot be inferred from context and thus remains too obscure to be assessed. Many of these lemmas are not even recorded in the main dictionaries. This formative element seems to be present in words such as *alwanzena-* 'enchanted', *araḫzena-* 'foreigner', *ḫerina-* (a word used in connection with the term for 'fire'), *kapina-* 'thread', ^{GIŠ}*karpina-* 'a (type of) tree', ^{LÚ}*kireštenna-* 'priest', *lappina-* 'firestarter, wick, tinderbox', *lappina-*$^{(SAR)}$, a phytonym indicating 'a (type of) garden plant', etc. Other entries are attested, perhaps segmentable differently, with a suffix of the type *-šīna-/šēna-*, which, according to Melchert's opinion (2002), would represent "covert compounds" rather than derivatives ending in *-ena-/-ina*.

In the case of $^{(GIŠ)}$*kalmišina-/kalmišana-* 'burning log', for example, if one follows Melchert's interpretation, it would be a "hidden" compound in *-šīna-/-šēna-* (cf. Pozza 2023 for the problems connected with graphic variants with *-a-* vocalism). Despite the lack of sufficient etymological evidence, that *kalwišina-*$^{(SAR)}$ 'edible plant or vegetable' possibly could be interpreted as a "hidden" compound of the type **kalwi* + *šīna-/šēna-*. There are

[16] This led Starke (1990: 473 ff.) to suggest a Luwian origin for the form.

[17] According to Rieken (1999: 256), an original **h$_1$y-è(h$_2$)-teh$_2$* should be postulated, whose final [-ắda] was interpreted by speakers as a variant of [**ắdar] lacking /r/ (see the text for full discussion).

[18] This suffix does not appear listed by Hoffner & Melchert (2008: 53ff. and 2024: 87ff.) among the main suffixes forming derived stems from nouns and adjectives.

cases such as *parḫuena-/parḫuina-* (one of the most frequently cited ingredients of magical material in Hittite ritual texts), whose formal and semantic interpretation is far from straightforward: 'oats' according to Francia (2020); 'a sweet beverage or vegetable product' according to Kronasser (EHS: 183); 'fermentation matter of cereals', 'material for beer production' according to Puhvel (HED PA: 122ff.); 'a type of grain' according to Tischler (HEG P: 457). Puhvel (2009, 2011: 72; HED PA: 122) considers it a lexeme of Indo-European origin, linking it to Greek φρείατα, Latin *ferv*(*e*)*ō* 'to boil', ferment 'yeast, malt liquor', Old English *brēowan* 'beer' (< **bhér-E₂-*, **bhr-éE₂-*(*w*-) 'to pant, boil, ferment', **bhr*(*e*)*Hwr/n-* 'rising, fermentation', genitive **bhr̥Hwén*(*o*)*s*)[19]: the meaning to be attributed to the word should therefore be 'fermentation beverage'. The prehistoric genitive of *parḫuena-/parḫuina-*, **bhr̥Hwéns* (realised as *parḫuenaš*), to which Puhvel assigns the value of '(beverage) of fermentation', would need to be reinterpreted as *parḫuena-* in the manner of $^{(LÚ)}$*kururaš* '(man) of hostility' > ^{LÚ}*kurura-* 'enemy' (see Yakubovich 2006) and *paḫḫuenaš* '(attack of) fire', genitive of *paḫḫur* 'fire'. Conversely, Francia (2020), who discusses all the passages in which the word is documented, considers that *parḫuena-/parḫuina-* would not indicate a beverage but rather oats (which can also be used to produce beer), characterised by calming properties that act both on the nervous system and the gastrointestinal tract (*ibid*.: 136). Therefore, it is impossible to clearly identify a verbal root from which it could derive (in cases like these, according to Puhvel, 2009: 77, one would be dealing with "hidden Anatolian derivates of otherwise common primary verbal roots"). As illustrated by the discussion, there is no certainty regarding the meaning conveyed by the lemma (still under debate) or whether it might be a word of Indo-European origin.

Among the words ending in *-ena-/-ina-*, some have obscure meanings and unknown or uncertain etymology (cf. $^{(LÚ)}$*ḫamina-/ḫamena-, ḫarmina-*, etc.), and others very likely represent loanwords (adapted from the replica language, cf. *ḫerina-* 'cedar wood; oven', possibly from Sumerian EREN 'cedar' → Akkadian *erēnu* 'wooden instrument'; *kulina-*, probably a Hurrian attribute referring to Ishtar; *kurupšini-*/$^{(LÚ)}$*kurupzina-* qualifier of material/shape of rhyta, perhaps related to Akkadian *kupuršin*(*nu*)*m*, a qualifier of the word for 'gold'), and still others, of probable Indo-European origin, interpretable as deverbal formations (cf. Gusmani 1978, Puhvel 2009) formed by adding the Indo-European suffix **-ĭ̄no-*[20] (possibly ^{GIŠ}*karpina-* 'a (type of) tree' < **(s)kerp-* and *lappina-* 'wick' < **leh₂p-*).

From these brief observations, one can clearly understand how complex it is to reason about the possible productivity of the suffix *-ena-/-ina-*, whose status and origin remain uncertain. The scarcity of attestations (together with the fact that many lemmas ending in *-ena-/-ina-* are, in fact, *hapax legomena*) and the strongly multilingual context in which the Hittite language is documented pose problems when providing general remarks regarding derivational morphology. The issue is further complicated, as shown, by the absence of a credible derivational base and a more than uncertain etymology.

[19] With **E₂* Puhvel (HED A: x) means "[a] voiced e-coloring laryngeal, Hittite *ḫ-*, *-ḫ-*".

[20] Or **-eyno-*? Or **-i(H)n(e)h₂-*? For the details, cf. Pozza (2023).

2.2. The case of *-i-* and *-u-* adjectives

In Hittite, as in other ancient Indo-European languages, *-i-* and *-u-*adjectives are quite common. The suffix *-i-* is not productive (cf. Hoffner & Melchert, 2008: 54 ff.),[21] and the high frequency of *-i-*nouns and adjectives is due to numerous loanwords from Hurrian formations in *-i-* (cf. Berman 1972: 9). Primary (underived) adjectives in *-i* include words such as *kappi-* 'small', *nakkī-* 'heavy; difficult', etc. Among the underived *-u-*adjectives are words like *idālu-* 'bad, evil', *tepu-* 'small', *panku-* 'whole', etc.

It is not always easy to determine the relationship between a given adjective in *-i-* or *-u-* and its possible base form, whether verbal or nominal, when the latter is not attested in the available documentation. Likewise, it is not entirely obvious whether the adjective should be considered prior to the correlated verb or *vice versa*, as in the case of Hitt. *šuu-* 'full' (cf. Alfieri & Pozza 2024: 162-163). Berman (1972: 188–189), Watkins (1975: 378), and Weitenberg (1984: 136), for instance, argue that *šuu-* derives from the verb *šuwa(i)-* 'to fill, to be full' by means of the addition of a suffix *-u-*. However, the prevailing view is that the verb should be interpreted as de-adjectival (HEG Š²: 1128; 1219: "sicherlich denominativum"; EDHIL: 797). If this latter proposal is accepted, *šuu-* should be regarded as a primary adjective,[22] even though, diachronically, it can be traced back to the (verbal) root *sew-(H)-* 'to fill' (HED ŠE/ŠI/ŠU: 134; LIV²: 539, s.v. *seu̯h₃-* 'to be/become full', "nur anatolisch").

A similar situation is found with *daššu-* 'strong; heavy; difficult' (possibly connected – cf. Kellogg 1925: 28 – with Gr. δασύς 'dense, thick' and Lat. *dēnsus* 'dense', or – cf. Juret 1941: 51 – with Skr. *dáṁsas* 'miraculous power'). It may represent a *-u-* deverbal adjective from *dašš-* (as argued by Kloekhorst, EDHIL: 854), although the base *dašš-* is not itself attested (the causative *daššanu-* 'to fortify' is documented, however, parallelling *tepnu-* 'to diminish', from *tepu-* 'small, little', itself derived from a non-attested base *tep-*). Alternatively, because no base form is documented, *daššu-* could be regarded as a primary adjective. Further evidence that the scarcity of documentation in Hittite sometimes prevents us from making clear-cut decisions about whether a given form is derived from its base is provided by *šarku-* 'eminent, illustrious, powerful'. Its derivation from the verbal base *šark-* 'to ascend, rise' (at present only attested in the iterative *šarkiške/a-* 'to be good', EDHIL: 734) is taken for granted by Gusmani (1968: 94). However, as noted by Kloekhorst (EDHIL: 734), the meaning 'to ascend' attributed to the base *šark-* in such an authoritative source as the CHD (Š: 268) ultimately rests (as acknowledged by the editors themselves) solely on the presumed connection with the adjective *šarku-* and on its semantic nuance associated with the concept of 'height', despite the absence of contexts in which the verbal meaning is incontrovertible.

[21] For the so-called "*-i-*mutation" of Luwian see in particular Oettinger (1987), Starke (1990), and Rieken (1994, 1999, 2005).

[22] The form *šuwant-* should be interpreted, according to Oettinger (1979: 296), as an ancient *-nt-* extension of *šuu-*, later reanalysed as the participle of *šuwa(i)-* 'to fill'.

In the same way, from a formal perspective, the adjective *kappi-* 'small' can be considered either as a primary adjective, or as formed from a verb such as *kapp(ai)-* 'to diminish, reduce' (EDHIL: 439; HED K: 62), which, however, is only inferred from the participle *kappant-* 'small'. Moreover, as mentioned earlier, the suffix *-i-*is not generally productive in Hittite. One may therefore argue, with Hoffner & Melchert (2008: 52–53), that both the adjective *kappi-* and the formation *kappant-* are preserved in the lexicon as primary words, being ultimately derived diachronically from the unattested root *kapp-* (on *kappant-* cf. also Dardano 2007: 13-14). In turn, the formation *kappant-* lends itself to a further interpretation, namely that of Frotscher, who considers it de-adjectival.

> Die Bildung *kappă̆nt-* 'klein' ← *kappi- / kappaḭ-* spricht in der Tat dafür. Es findet sich nämlich kein Verb **kappae-zi* 'verkleinern', wozu dieses Ptz. sein könnte. Stattdessen ist es einfacher *kappānt-* aus **kapp-aḭ-ant-* ← *kappi- / kappaḭ-* 'klein' herzuleiten und als denominale *-(a)nt-*Bildung zu werten (Frotscher 2013: 40).

However, as also emphasised by Frotscher, unlike the *-u-*ablaut stems (which exhibit derivations such as *idalawant-* ← the stem *idalaw-* from *idālu-*), for the *-i-* ablaut stems there do not seem to be documented derived formations from the *-ay-* stem (and thus, potentially, *-aiy-ant-* > *-ānt-*) that could support this hypothesis: *šuppiiyant-* ← *šuppi-* 'pure', which appears to go against the trend, is a non-ablaut *-i-* stem and therefore not conclusive. Frotscher (*ibid*.: 163) does not exclude verbal derivation, thereby illustrating the objective difficulty in taking a definitive stance on the matter.

A similar case is that of *ḫarki-* 'white, bright', which appears to be associated with a verb attested with different orthographic variants:[23] *ḫar-ki-eš-zi* (KBo 2 i 44-45), *ḫar-ki-e-eš-zi* (KUB 15.39 + 12.59 ii 16), *ḫar-ki-i-eš-zi* (dupl. KBo 39.8 + iii 4), and *ḫar-ki-i-iš-zi* (KUB 27.67 ii 28). These spellings allow for readings such as *ḫark(i)ešš-* (HED Ḫ: 170), *ḫarkiyešš-* (EDHIL 307), and *ḫarkešš-* (HEG A-K: 177); the meaning is the same ('to become white'), but the synchronic morphological interpretation of each form differs: *ḫarkiyešš-* is a denominal verb derived from *ḫarki-*, parallel to *tepaw-ešš-* from *tepu-* (de-adjectival verbs are constructed on different ablaut grades); *ḫarkešš-* is a case of suffix substitution derivation based on *ḫarki-*, parallel to *tepnu-* from *tepu-*; and *ḫark(i)ešš-* would be compatible with both perspectives. However, based on graphic considerations of the variants, it is more likely that the spelling <-ki-i-> alludes to the denominal *ḫarkiyešš-* and that the suffix therefore followed (and did not replace) the /i/ of the adjectival stem. This would exclude a suffix substitution derivation, which, as shown (cf. § 2), involves a diachronic etymological interpretation but does not imply morphological productivity synchronically. Furthermore, there are no attested *-i-*adjectives built on athematic verbal bases through derivation by addition and, consequently, *ḫarki-* can be interpreted as a primary adjective without internal structure.[24]

[23] Much has been written on *scriptio plena*. We refer especially to Kloekhorst (2014).

[24] The relationship between graphic considerations of the variants and morphological productivity may not immediately reveal a clear causal connection. Nevertheless, the reading deemed most appropriate appears to reflect a derivation by addition (rather than substitution) from the simple adjective.

2.3. The case of *ḫūmant-* 'each, all, entire, *ḫappina(nt)-* 'rich', *ḫaršallant-* 'angry, furious'

The suffix -(*a*)*nt-* conveys more than one function in Hittite (cf. § 1, as well as Hoffner & Melchert 2008: 55-56; 2024: 89-90): it forms participles (*miyant-* 'in bloom' < *mai-/mi-* 'to grow, to prosper'); denominal possessive adjectives (*perunant-* 'rocky' < *peruna-* 'rock'); has an "individualising" function (*ḫamešḫant-* 'the (upcoming) spring' < *ḫamešḫa-* 'spring'), alternates – without semantic variation – with forms that lack it (*irmala-/irmalant-* 'sick'); and, in some cases, reflects the addition of *-t-* to a stem in -(*a*)*n-* (as in *išpant-* 'night' compared, for example, to Av. *xšapan*).[25]

Among the examples of lexemes that exhibit traces of internal structure without it being possible to demonstrate that they result from productive word-formation rules – especially because none of the hypothetical base forms of these derivations are attested as such (hence the presence of the asterisk) – are the cases[26] of *ḫūmant-* 'each, all, entire' and *ḫappina*(*nt*)- 'rich' (on primary adjectives in **-nt-* in Anatolian, the important study by Dardano 2007 is of notable relevance). Even in this case, the bases **ḫū*(*m*)- and **ḫāpp*(*in*)- are not documented[27], or at least not with the required value, as in the case of *ḫāpp-*, which, from a purely synchronic point of view, means 'to unite, to attach (used impersonally or in the middle voice), to manage, to work'.

Therefore, the meaning 'to abound, to be rich,' which is the prerequisite to justify a derivative *ḫappina*(*nt*)- 'rich,' can be ascribed to *ḫāpp-* only on etymological grounds (that is, if the connection with Sanskrit *ápnas-* 'possessions,' Latin *ops* 'wealth' is accepted), but from a purely synchronic perspective, none of these meanings can serve as the basis for constructing an adjective with the value 'rich'.[28] That *ḫappina*(*nt*)- could be correlated with *ḫāppar-/ḫāppir-* 'trade, business', based on an original heteroclitic "Proto-Hittite" form such

[25] See Dardano (2010: 6), who states that «the presence of the same suffix in both participles, more properly defined as verbal adjectives, and in primary adjectives should not be surprising», when compared with the historical continuations of the suffixes **-lo-*, **-no-*, and **-to-*; «the common denominator is the use of a deverbal derivational suffix (originally from roots, only later deverbal) also in denominal formations» (*ibid*.: 7). Consider, for example, *ḫattant-* 'intelligent, sharp', which, as noted by Dardano (2007: 17, note 61), is recorded as an autonomous lexeme by Puhvel (HED Ḫ: 260-263), separately from the participle *ḫattant-*, from < *ḫat*(*t*)-, *ḫatta-* 'to pierce, to strike', with the meaning 'pierced, struck'.

[26] On *ḫappina(nt)-* and *ḫūmant-* see the discussions in HW^2 (Ḫ: 231-232, ḪE-ḪU: 712ff.).

[27] Clearly, the unattested forms in Hittite texts could appear in other lost texts, or related forms might be found in other Anatolian languages. For the dissimilation of /w/ to /m/ before /u/ (in the case of postulating a suffix *-want-* on the base **ḫū-*, as an alternative to *-ant-* on the base **ḫūm-*), see Melchert (1994: 109; 127). For this matter, also consult HEG (Ḫ: 381).

[28] Laroche (1963: 72) translates an occurrence of the verb as "avoir en abondance" (*takkuš-maš UL-ma ḫapzi ta natta ḫazzianzi* "mais s'ils sont dans le dénuement [lett. 's'il n'y a pas pour eux abondance'], on abat un porc", KBo 11.34 i 5). This translation is rejected by more recent studies: Neu (1968: 45, fn. 1; 1974: 83) translates "wenn es sich ihnen aber nicht fügt, stechen sie (es) nicht ab" ["but if it does not fit them, they do not stab (it)"]; Puhvel (HED Ḫ: 251) translates "but if it does not work out for them, they do not stick [it]", and the entire etymology is rejected by Kronasser, who considers *ḫappina-* "ohne nachweisbares Grundwort" ["without an attested base word"] (EHS: 182). See also HW^2 Ḫ: 196, s.v. *ḫap(p)-* 'sich fügen': "kein Zusammenhang besteht mit *ḫappina(nt)-* 'reich', *ḫappar-* und *ḫappira-* 'Stadt (Dorf)'".

as **ḫapér/*ḫapén-*, is the view of Oettinger (1979: 353, 1981: 149),[29] but, even in this case, the etymological hypothesis does not change the fact that, from a synchronic point of view, it is a primary adjective.

That *ḫūmant-*, for example, could – according to Kimball (2007: 201 ff.) – represent the participle of an athematic compound verb of the type **h₂u-h₁em-/h₂u-h₁m-*, where **h₂u-* 'together' would constitute the preverbal element and **h₁em-* 'to take' the root (hence the meaning 'taken together' for the participle *ḫūmant-*), certainly cannot be ruled out (also supported by Frotscher 2013: 143, albeit with different arguments), from the point of view of a possible etymological interpretation. However, synchronically, deciding on the type of base represented by **ḫum-* solely based on its occurrence with the suffix *-ant-* is equally (if not more) questionable than considering *ḫūmant-* as stored in the lexicon. "Errstarrte Bildung" also for Frotscher (ivi: 144), exited early and subsequently became isolated from the participial system.

Semitransparent words that – although marked by some internal structuring – cannot be traced back to others through derivational rules can thus be considered "lexicalised"[30], because they are stored in the lexicon: *ḫappina*(*nt*)- 'rich', *ḫūmant-* 'each, all', and *tepu-* 'small', for example, cannot be traced back to any attested Hittite base and are therefore preserved in the lexicon, even though they exhibit the suffixes -(*a*)*nt-* and *-u-*. The suffixes *-u-* and -(*a*)*nt-* are "real" morphemes in words like *ḫuišu-* 'alive' < *ḫuiš-* 'to live' and *karšant-* 'cut' < *karš-* 'to cut', but they are "quasi-morphemes" in *tepu-* and *ḫappina*(*nt*)-, thus semi-frozen morphemes formally identifiable but not productive functionally (Aronoff 1976: 11), similar to the sequences *-u-* and -(*a*)*nt-* in *idālu-* and *ḫūmant-*.[31]

The complexity of the univocal classification of certain forms has been the subject of study by Dardano, who, in the already cited 2007 article (*ibid*.: 16-17), mentioned some entries of ambiguous interpretation, including *enant-* 'tamed' and *parrant-*, of uncertain meaning, used in reference to straw, for which the classification tends to oscillate between participle and adjective. Dardano notes that the former is lemmatised as *enant-* by Puhvel (HED E/I: 271) – who traces it back to a (unattested) root *en-* < **ain-* 'to agree' – which Tischler (HEG A-K: 106) derives from *annanu-* 'to instruct, to teach' (of unclear etymology), and it is not classified explicitly by HW² (E: 37). For *parrant-*, more complex in semantic reconstruction, dictionaries (CHD P: 135, HEG P: 441) fluctuate between classifying it as adjective or participle, leaving the question open. The same applies to other forms discussed by Dardano, among which are *tatrant-* 'pointed, sharp; aggressive' and *paprant-* 'impure', for

[29] Cf. also Rieken (1999: 315).

[30] Cf. Bauer (2001: 27): "[…] we can note that while lexicalisation as discussed just above is a process which affects individual words diachronically, the result is that at any synchronic moment different words will be at different stages of lexicalisation, the diachronic process being reflected in the synchronic status of individual words."

[31] Another example illustrating the difficulty in choosing between a verbal or nominal derivation is that of *išḫaškant-* 'bloodstained', which could be a participle from **ešḫar-šk-* with **-ršk-* > **-šk-* (cf. the *hapax eš-ḫa-ri-eš-ki-it-du* in EHS: 456, 491, 506 and HEG A-K: 115), or a syncopated participle of the iterative verb *ešḫaneški-* derived from **ešḫaniya-* 'to blood': **ešḫan(i)škant-* > *ešḫaškant-/išḫaškant-* (HED E/I: 309). Both verbs, however (although their base forms are reconstructed but unattested), can be traced back to the noun *ešḫar* 'blood'. See also the discussion in EDHIL: 258-260.

which no base verb is documented, only the corresponding causative formation (*tatraḫḫ-* 'to incite, provoke', *papraḫḫ-* 'to make impure'). Dardano (2007: 22) concludes that most of the *-(a)nt-* formations in Hittite, unlike the corresponding ones in other Indo-European languages, should be considered as verbal adjectives (primary, root-based), and therefore belong to the realm of derivational morphology, not inflectional (as participles do). It is essential to observe that the choice to consider an adjective as primary, not directly traceable to an unattested verbal form synchronically, does not relate as much to its evaluation in light of the Caland system, and thus to whether a form (if root-based and not deverbal) can fit into the system. What is under discussion here – following Dardano's line of reasoning – is that, in the absence of an attested base, assumed solely on the formal structure of the derivative and the (diachronic) reconstruction of its etymological basis, a derived lexeme, although morphologically transparent in its internal structure, should be considered as primary in the speaker's lexicon (that is, in their competence).

Finally, the case of *ḫaršallant-* 'angry, furious' is noted, perhaps a participle of an unattested (denominal?) **ḫaršal(l)a(i)-* (HED H: 186, HW^2 Ḫ: 341) or (HED H: 186) a denominal derivative in *-nt-* from a **ḫaršalla-* (the relation to *ḫaršar/ḫaršan-* 'head' should not be excluded but only diachronically, as a result of dissimilation from an original **ḫaršan-ant-*?). "Ohne Grundwort" in Kronasser's opinion (EHS: 266). Tischler (HEG A-K: 183) cites the possible derivation from the verb *ḫarš-* 'to tear, to break'. Nevertheless, even in this case, as emphasised by Kronasser (EHS: 266), "wenn jedoch Grundwörter fehlen, läßt sich oft nicht einmal die ursprüngliche Wortart mit einiger Sicherheit feststellen, da Partizipia zu Nomina und Adjektiva zu Substantiva werden können". A primary formation, therefore, even if endowed with a semi-transparent internal structure.

2.4. The case of *armawuant-* 'pregnant', *ešḫarwant-* 'bloodstained; blood-coloured', *ḫuišwant-* 'alive', *innara/uwant-* 'vigorous, strong', *le/iliwant-* 'quick, winged, urgent', *mišriwant-* 'luminous, brilliant'

Let us now examine some cases of formations in *-want-*, a suffix which, as has been seen (§ 1), attaches to both nouns and verbs. Even in this case, determining whether the formation is nominal or verbal is not straightforward, and it cannot be excluded, in the absence of the derivational base, that it was a primary lexeme stored in the lexicon (see also § 1 and fn. 12). Fundamental to the analysis and evaluation of the suffix *-want-* are the monograph by Maier (2013) and Frotscher's doctoral thesis (2013: 41 ff.), in which the latter classifies the deverbal formations (participles) and the nominal constructs in *-want-*, categories between which it is not always easy to draw a clear distinction, especially when both verbal and nominal bases are present. The merit of this study lies in the extensive discussion (see also below) of the substitutional suffixation process based on the derivational model of the type *-ant-* ~ *-aḫḫ-, -ātar-, -ē-, -ēšš-* (exemplary, from this point of view, is the final schematisation, *ibid.*: 344-353).

Of difficult resolution, for example, is the case of *armawant-* 'pregnant', deverbal from *armai-* 'to be pregnant' (HED A: 157), as well as potentially nominal (EHS: 266), derived from *arma-* 'moon; month' (documented only in the Sumerogram ^{d}EN.ZU, ITU(KAM), and

the Akkadogram ^{d}SÎN), or from *arma-* 'pregnancy' (HEG A-K: 62). In Maier's view (2013: 20-21), in addition to the deverbal hypothesis (which would have first presupposed derivation from the genitive *armawaš* of the unattested verbal noun *armawar*), a derivation from the *-aw-* stem of a *armu-* could also be outlined. In this case, however, in the absence of attestation of the base noun, it seems preferable to lean toward the first hypothesis (also supported by Frotscher 2013: 89-90, 99).

The adjective *ešḫarwant-* 'bloodstained, blood-coloured',[32] again, can be interpreted as a denominal in *-want-* derived from *ešḫar-* 'blood', but at the same time it may represent the participle of a *išḫarwai-*, itself a denominal verb built on the unattested noun *išḫaru-* 'bloodyness', connected to *ešḫar-* 'blood'. The first option is based on actually documented data. It therefore appears preferable, because the verbal base *išḫarwai-* and the noun *išḫaru-*[33] are unattested (Oettinger 1988: 284, Hoffner & Melchert 2008: 61). For a more detailed discussion of the various proposals, see Otten & Souček (1969: 53), HEG (A-K: 115), Rieken (1999: 483 ff.), Maier (2013: 61-63).

Furthermore, *ḫuišwant-* 'living, alive' could be interpreted[34] as a deverbal adjective in *-want-* from *ḫuiš-* 'to live', as an extension in *-nt-* of *ḫuišu-* 'alive', or as a participle in *-(a)nt-* built on *ḫuišwai-* 'to be alive', a denominal verb derived from *huišu-* 'fresh, raw', which in turn is a deverbal adjective built on the verbal base *ḫuiš-* 'to live' (cf. EHS: 267, HEG A-K: 268).

A helpful device to orientate oneself toward the participial interpretation is to evaluate the position relative to the possible noun to which it refers, because participles follow the noun in all uses in Hittite. In contrast, adjectives tend to precede the noun in attributive function and follow it in predicative function (Francia 2001). However, these are tendencies, not strict rules, and attributive adjectives in -want- can quite freely precede or follow the noun (Hoffner & Melchert 2008: 272). Therefore, distinguishing and thus choosing between the function of a predicate and that of an adjective can be difficult, especially in cases of an adjectival predicate without a copula.[35] The proposal of Frotscher (2013: 136, 202-204) and Maier (2013: 47) is that this is a participle in *-ant-* (of the stative verb *ḫuišwai-* 'to be alive') functioning adjectivally, with the syntactic behaviour of an adjective (attributive, predicative, substantivised, and – albeit rarely – even adverbial).[36] Regarding the participles of stative verbs, Frotscher (2013: 203) observes that "the participle expresses a state that has

[32] Cf. for example (KBo 17.1 i 24-25) *weššanda=ma išḫarwantuš TÚGḪI.A-uš* "they wear blood-red garments" (HED E/I: 311). On *ešḫarwant-* and *ešḫarnuwant-* cf. Frotscher (2013: 41).

[33] "On the basis of *išḫarwant-*, however, the stem *išḫaru-* received some productivity, resulting in forms such as *ēšḫarwaḫḫ-*, *išḫarwieške/a-* […], and *išḫarwīl*" (EDHIL: 260).

[34] See the various hypotheses proposed in the synthesis of Maier (2013: 45-47).

[35] Consider the case, discussed in Alfieri & Pozza (2024): *ḫalkiaš ḫaršār išḫiy-and-[a]* [Z]ÍZ$^{ḪI.A}$*-aššš=[a] ḫaršār išḫiy-and-a* (KBo 17.1 iv 19-20, Otten & Souček 1969: 37 'die 'Köpfe' von Gerste (sind) (zusammen-) gebunden, und die 'Köpfe' von Spelt (sind) (zusammen) gebunden'). Otten & Souček translate *išḫiyant-* as '(are) bound' (the text lacks the copula). The passage, however, can be interpreted either as "the barley ears are bound together" or "the barley ears bound together."

[36] For example, for the participle *ašant-* in the adjectival value of 'true' (in attributive and predicative function), see the detailed analysis by Cotticelli-Kurras (1991: 158 ff.).

become a property, whereas the finite form does not express a property, but only the state. The participle is therefore an adjective".

If, on the other hand, the base form of adjectives in -(*a*)*nt*- or -*want*- is not attested, it would be appropriate, according to the same principle, to consider these adjectives as non-derived, albeit endowed with a semi-transparent internal structure. Indeed, although the formation of adjectives from nouns is fairly productive in Hittite, in many cases the nominal bases from which certain adjectives might appear to be derived are not attested: in addition to the examples already mentioned, this is the case of words such as *innara/uwant*- 'vigorous, strong', *le/iliwant*- 'fast',[37] *mišriwant*- 'bright, shining', respectively from the unattested **innaru/a*-, **le/ili*-, and **mišri*-. From a strictly synchronic point of view, we should consider them primary formations, despite their later attestation in other derived forms such as *mišriwaḫḫ*- 'to make bright, brilliant', *mišriwatar* 'brightness', *mišriwešš*- 'to become bright', *leliwaḫḫ*- 'to hurry', *in*(*n*)*ara*(*wa*)*ḫḫ*- 'to strengthen', *innarawawar* 'strength', *innaraweššš*- 'to become strong', etc.

As for *innarawant*-, for example, Weitenberg (1984: 189) reconstructs the base **innaru*-, while Frotscher (2013: 54), underlying that "**innaru*- is, however, not attested as such", states that "instead, *innaraṷant*- is a -*ṷant*-adjective derived from **innara*-, as it appears in the adverb (< Nom.-Acc.Pl.n.) *innarā* 'intentionally, diligently'". The same view had been expressed some time earlier by Melchert (1984: 80), who also suggested that the verbal derivatives *innaraweššš*- and *innarawatar* were formed on a base **innarawa*- extracted from *innarawant*-, following the model of pairs like *pittalwa-/pittalwant*-, etc. Finally, it is also worth mentioning Maier's (2013: 56-59) detailed synthesis, according to which one could reconstruct an abstract nominal base **innara*-, built on **innar*- (according to Hrozný 1917, **innara*- could represent the outcome of a compound such as **h₁en-h₂nor*- 'endowed with internal vigor'; *contra*, EDHIL: 387).

Etymologically, then, **mišri*- 'glitter' in *mišriwant*- can be traced back to **miš*- 'to sparkle' (< **meys*- 'to shine') plus the suffix -*ri*- (Hoffner & Melchert 2008: 59; 2024: 93). Nevertheless, the etymology is uncertain (EDHIL: 582). From a morphological standpoint, according to the same authors (2024: 90), "*mišriwant*- was reinterpreted as *mišriwa-ant*- like *pittalwant*- and -*ant*- was deleted in the formation of *mišriw-aḫḫ*- 'to make splendid, perfect'". According to Neumann (1962: 155), *mišriwatar* and *mišriwešš*- would instead derive respectively, "mit stärker Syncope" from **mišri-want*-(*a*)*tar* and **mišri-want-eš*. Forms like *mišri-want*- would then have been reinterpreted as *mišriw-ant*- before -*ant*- was deleted in the formation of verbs such as *mišriwaḫḫ*- 'to make splendid, perfect', following an extension of the substitution pattern found in -*i*- and *u*-stems, and -(*a*)*nt*- (although suffixation by addition, as already noted, should have been the regular pattern in Hittite; cf. *ḫatku*- 'tight' > *ḫatku-ešš*- 'to become tight', etc.). Along the same lines (elision of °-(*a*)*nt*-) see Oettinger (1979: 241). Kronasser (EHS: 401), opposed to the hypothesis proposed by Neumann (cf. above), does not exclude the reconstruction of a base **mišriwa*- and motivates

[37] For a thorough interpretation of the entry, see Frotscher (2013: 83-84), who does not exclude a verbal derivation from **lelai-i/*leli*-, with the presumed meaning 'to move quickly' (for the issues related to the originally transitive semantics of the verb, I refer directly to Frotscher's work). See also Maier (2013: 94-95).

the derivation of *mišriwešš-* from *mišriwant-* on the basis of a proportional analogy of the type *idalawant- : idalawešš- = mišriwant-* : x. All these hypotheses, although plausible, do not resolve the problem of the unattested synchronic base (**mišri-*? **mišriwa-*?).[38]

That speakers may have carried out false segmentations of the forms in *-want-*, following the obsolescence of the original bases **le/ili*, **mišri-*, and **innaru/a-* (the latter, in the *-a*-stem, inferred from the adverb *innarā*, but never attested autonomously as a noun), and that consequently they originated the above-mentioned verbal forms from bases such as *mišriw-*, *innaraw-*, and *leliw-* (cf. *supra*) is certainly a plausible hypothesis. However, the fact that a nominal base is not (any longer) documented autonomously but only indirectly through multiple derivatives does not allow us to classify such formations – at a synchronic level – as denominal, but rather as already lexicalised forms. Additionally, false segmentations or back-formations, which generalise a model of morphological relation, are at best semi-productive sporadic processes.[39]

Even in cases such as those just mentioned, essentially, despite formations with a (semi)-transparent internal structure, we would be dealing, from a synchronic perspective, with "primary" adjectives. Frotscher (2013: 54 ff.), however, considering the verbal derivatives of the forms discussed above as formed through suffix substitution – following the ideas previously proposed by Neumann (1962: 154-155), Oettinger (1979: 240 ff.) and Hoffner & Melchert (2008: 51; 2024: 85) – argues that the base for their formation was not a noun (which, as we have seen, is not attested), but that the derivational formation model started precisely from the stem in *-want-*, according to a derivational pathway that originated from the substitution of the participial suffix with that of the factitive verbs, from which (-)*ant-* → (-)*aḫḫ-*, (-)*ešš-*, etc.

Regarding the factitive formations in *-ešš-*, Frotscher envisions two developmental models, as in the case of an adjective like *parkui-/pargaw-* 'pure', which shows a dual outcome in the factitive derivatives: the form *pargaw-ēšš-* would derive from the (unattested) adjective in -(*a*)*nt-* **pargaw-ant-* (fully consistent with what was previously observed for the factitives in *-aḫḫ-*), whereas *parku-ēšš-* would have as its base the adjectival stem without the *-i*-extension. As can be seen, the perspective adopted by Frotscher, while entirely reasonable, differs (though not contradicts) the line pursued here: postulating a historical origin of *pargaw-ēšš-* from a hypothetical **pargawant-* is a valid diachronic-reconstructive operation, but in fact it does not conflict with the idea that both factitive formations derive from the simple adjectival base and that, for speakers, the productive model was not **pargawant-* → *pargaw-ēšš-*, but rather *pargaw-* → *pargaw-ēšš-*.

[38] More recently, see the detailed discussion in Maier (2013: 104-106). See also Frotscher (2013: 54-55, 348). Cf. Tischler (HEG L-M: 217) and Puhvel (HED M: 164) for further bibliographical references.

[39] Cf., among others, Matthews (1991: 69).

3. Final remarks

As has been observed, when attempting a synchronic morphological segmentation of certain Hittite adjectives containing productive suffixes, it is not always possible to identify the base form (verbal, nominal, or adjectival) from which the derivative should come. Consequently, this does not allow confirmation that a given suffix attaches exclusively or predominantly to a specific type of base (precisely because it is not attested). Indeed, derived Hittite adjectives can be classified as denominal (e.g., *genzuwala-* 'kind-hearted, merciful' < *genzu-* 'mercy'), deadjectival (e.g., *appezzi*(*ya*)- 'posterior, subsequent' < *āppa* 'behind, after'), deverbal (cf. participles like *kariyant-* 'covered' < *kariya-* 'to cover', or forms like *parku-* 'high' < *park*(*iya*)- 'to increase, elevate, (make) grow', *tarḫuili-* 'strong, powerful' < *tarḫu-* 'to prevail, conquer'), deadjectival with extended suffixes (cf. forms in -(*a*)*nt-* like *dannarant-* 'empty, smooth' compared to *dannara-* id., or *arawanni-* 'free' compared to *arawa-* id.), as well as lexicalised compounds (cf. above, *dāyuga-* 'of two years') or of heterogeneous nature (*kurur* 'hostile, enemy'),[40] *šanezzi-*/*šanizzi-* 'pleasant, excellent, valuable[41] etc.).

Hoffner & Melchert (2008: 52; 2024: 86) clarify that nominal or adjectival formations in Hittite, for which it is not possible to identify a base form of derivation (regardless of their status in the pre-Hittite phase), should be considered as underived: therefore, in such cases, from a methodological point of view, these forms should be regarded as lexemes stored in the lexical competence of speakers as unitary lexemes (even though their internal structure is transparent and it is possible, diachronically, to decompose their constituents). Consider the case of *šuppišduwara-* 'brilliant,' perhaps derived from *šuppi-* 'pure, purified' + **išduwara-*, an unattested verbal noun, in turn derived from *išduwa-* 'to be manifest, to be revealed' (Neu 1970: 69), but whose connection with *šuppi-* remains somewhat uncertain (cf. EDHIL: 791). In fact, the semantic connection between the two composing elements of the compound is problematic (cf. Melchert 2017b: 179).[42]

It is also evident, as already mentioned, that the lack of attestation of a base form, in the case of a language such as Hittite, can be accidental, and therefore, the evaluation of a form as "primary" with respect to the typology of synchronic segmentation offered solely by the documented forms is not the most correct solution. However, wishing to distinguish between forms objectively derived from attested bases and forms whose synchronic derivation is uncertain (also in light of, as seen, the non-unique attachment of a derivational morpheme to a specific base), the approach taken in the examination presented here has been to opt for classifications that, as objectively as possible, are founded on the currently available data.

[40] On its (secondary) adjectival value, cf. Hoffner & Melchert (2008: 61, 117), EDHIL: 496, HED K: 278, HEG A-K: 665 and, most of all, Neu (1979).

[41] Perhaps derived from *šani-* 'the same' + suff. *-ezzi(ya)-*? (EDHIL: 723). According to Berman (1972: 201), obscure formation.

[42] "The evidence of the Hurro-Hittite Bilingual shows that *šuppištuwar-ant-* is another possessive adjective in *-ant-*, 'possessing bosses, protuberances, appliqués' from the noun *šuppištuwara/i-* 'boss, protuberance, appliqué' (also spelled once *išpištuwarāš* at KUB 42.64 Vo 2). We are dealing with a derivative of the PIE root **spei-* referring to various pointed objects: cf. English 'spit' or 'spire'" (Melchert 2017b: 179).

As already noted by Gusmani (1968: 95 ff.), sometimes in Hittite the possible relation to a verbal base is no longer demonstrable, because the base formation has been lost or has become unrecognisable, even if documented as a root formation in other historical Indo-European languages: this is the case of *panku-* 'all, entire,' which diacritically can be traced back to PIE **bhengĥ-* (cf. Sanskrit *baṃbhayate* 'strengthens, increases', *bahú-* 'much, wide', gr. παχύς 'thick'), but whose Hittite verbal base is absent.

The same observation applies to *tepu-* 'few, scarce', an *-u*-adjective whose base, **tep-*, is not documented. The derived verb *tepnu-* could represent the extended *-nu-* form of the base (unattested) verb underlying the adjective *tepu-*, or it could be a deadjectival verb derived from *tepu-*. This ambiguity means that the parallel with Sanskrit *dabhnóti* 'damages', where the verbal base is attested, should not be given much importance (Gusmani 1968: 96). The productive derivational processes in the synchronic domain should thus be distinguished from etymological analysis. Kronasser (EHS: 418-419) had already questioned the distinction between etymological and synchronic analysis when discussing the concept of root in Hittite. Morphological productivity types (cf. Bauer 2001: 25) correspond to quantitative, qualitative, synchronic, or diachronic criteria: some are based on existing words, others on potential words. This recalls the idea that, in the speaker's perception, "the less morphotactically transparent, the more storage" (Mayerthaler 1987: 45). It is clear that the use of Hittite dictionaries and the various etymological proposals presented from time to time represent an essential source for a more complete evaluation of problematic lexemes, also to avoid confusing general reflections on the cognitive aspect of speaker behaviour with those derived from a rigorous diachronic investigation.

The same observation was made starting from the more general groupings – in the reference grammars – of derivatives presenting the same suffix, regardless of whether they are the result of productive rules, from which some less obviously classifiable forms were extrapolated. As seen, for example, *ḫappina*(*nt*)- 'rich' (cf. *supra*) is usually considered an -(*a*)*nt*-derivative,[43] *idālu-* 'bad, evil' as an adjective in *-u-, dāyuga-* 'of two years' as a compound (built on **dā-* < PIE **dwoyo-* 'two' – cf. the adverb *dān* 'for the second time' – although **dā-* is not attested in the expected form, because no synchronic rule predicts the deletion of *-n- from dān*)[44] etc.

Such situations are more easily framed within a scale (cf. especially Bybee 1985: 81ff. and Mayerthaler 1987: 46) whose extremes are represented respectively by complete storage

[43] Probably, as indicated in Hoffner & Melchert (2024: 90), "the synonymy of the *-a-* stem and *-ant-* stem in cases like *marša-/maršant-* and *pittalwa/pittalwant-* 'plain' and instances where only the longer variant survived (e.g., *marlant-* 'foolish' to **marla-*) permitted speakers to reanalyze derivatives of the base adjective as belonging to the *-ant-* stem, leading to *ḫappinant-* 'rich' > *ḫappin-aḫḫ-* 'to make rich' and *ḫappin-ešš-* 'to become rich'". In Oettinger's opinion (1981: 148) *ḫappina-* should be interpreted as a back-formation on the more frequently attested *ḫappinant-*. On the semantics of *ḫappina-, ḫappinant-* cf. Cotticelli-Kurras (1998).

[44] It is also true that in many languages the forms of lexemes entering into compounds do not have autonomy outside the compound, which does not imply that such compounds are not formed by productive rules nor that those forms depend on rules applying only in the context of the compound. However, synchronically, the relation between the adverb (including ordinal numeral, cf. HEG T1: 89) *dān* and the first member of the compound *dā-* is not documented in other formations, just as the cardinal numeral *dā*, deduced (since Hrozný 1917) precisely from *dāyuga*, is not (yet) attested.

and complete processing through rules, a *continuum* along which there exist possible morphological spaces occupying an intermediate zone, outlining "compromise pockets" corresponding to morphological structures that are partially stored and partially processed (because they are less transparent morphosemantically and morphotactically). Thus, adjectives like *ḫupigawant-* 'veiled' (← *ḫupiga-* 'veil' + *-want-*) or *išpantiya-* 'nocturnal' (← *išpant-* 'night' + *-iya-*) can be considered as clearly rule-processed and are placed at one end of the categorial *continuum*, while forms such as *ḫuelpi-* 'fresh, young' or *pittalwa-* 'simple, pure', which are certainly primary, lie at the opposite end as they are undoubtedly stored in the lexicon. The "compromise pockets" could contain cases like the already mentioned *alwanze/ina-* 'magical, practicing witchcraft', *ḫappina*(*nt*)- 'rich', *ḫūmant-* 'every, all', *mišriwant-* 'brilliant' etc., whose derivational morphemes are semi-frozen (cf. § 2.4) and whose bases (***alwanza-*, ***kalwi-*, ***ḫappin-*, ***ḫūm-* and ***mišri-*), not synchronically attested, are not immediately identifiable and uniquely associable with a specific part of speech. It is undoubtedly true that diachronic analysis can prove decisive in most of the problematic cases mentioned so far, mainly because Hittite data, unlike those of other historical Indo-European languages, are in some ways more complex, both due to the fragmentary attestation of some forms, due to phenomena linked to linguistic interference and the multilingual – also graphical – geographical context, and due to the difficulty that is encountered more than once even in reading – and thus in the consequent morphological evaluation – of a specific (and sometimes unique)[45] attestation, etc.

Therefore, evaluating data from both synchronic and diachronic perspectives is all the more critical. Consider the well-known example in Italian of *giornalaio* ('newspaper seller') vs. *beccaio* ('seller of goat meat' → 'butcher'): the former follows an Italian derivational rule (*giornale* 'newspaper' + *-aio-*), while the latter has the same suffix only in a diachronic view, as it continues Latin *beccarius* 'butcher' but cannot be traced back to Italian *becco* (which has an entirely different meaning, 'beak'). The analysis for Hittite is less evident because the documentation does not allow for such a clear distinction, but this does not mean one should be unaware of it. Nonetheless, what has been attempted here is a synchronic morphological overview of some formations, first to assess their "distance" from lexicalisation or derivation, categories whose boundaries are notoriously blurred.[46] The analysis, conducted on a sample of individual derivational types in the formation of adjectives, provides a substantial methodological indication of the necessity of careful case-by-case examination, to avoid generalisations that would flatten different chronological levels (Indo-European derivation, Proto-Hittite derivation, Hittite derivation with rules operating in the speaker's competence).

[45] For the relationship between *hapax legomena* and productivity in ancient languages, cf. Sandell (2015: 34-35).

[46] Refer to the bibliography cited in footnotes 10, 11 and 12 and *passim* in the work. For the quantitative results related to the Hittite language (which seems to present a higher percentage of primary adjectival structures alongside a large number of participial formations or those secondarily derived from verbal roots), see Alfieri & Pozza (2024) and the specific theoretical typological framework discussed therein.

References

Alfieri, L. 2009. La categoria dell'aggettivo in vedico. *Archivio Glottologico Italiano* 94(1). 3-41.

Alfieri L. 2014. *Qualifying modifier encoding and adjectival typology*. in Simone, R. & Masini, F. (eds.), *Word classes: Nature, typology and representations*, 119-139. Amsterdam – Philadelphia: John Benjamins.

Alfieri L. 2016. The typological definition of the (apparently historical) notion of root. *Archivio Glottologico Italiano* 102(1). 129-169.

Alfieri, L. 2020. The lexicalization of the adjective as an innovative feature in the Indo-European family. *Poznań Studies in Contemporary Linguistics* 56(3). 379-412.

Alfieri, L. 2021. *Parts of speech comparative concepts and Indo-European linguistics*. In Alfieri, L. & Arcodia, G. F. & Ramat, P. (eds.), *Linguistic categories, language description and linguistic typology*, 313-366. Amsterdam – Philadelphia: John Benjamins.

Alfieri L. 2023. Review of: Malzahn, M. & Fellner, H. A. & Illés, Th-S. (eds.). 2022. *Zurück zur Wurzel: Struktur, Funktion und Semantik der Wurzel im Indogermanischen: Akten der 15. Fachtagung der Indogermanischen Gesellschaft von 13. bis 16. September 2016 im Wien, Reichert, Wiesbaden*. *Ricerche Linguistiche* 1. 245-272.

Alfieri, L. & Pozza, M. 2024. The Hittite adjective in comparative perspective. *Archivio Glottologico Italiano* 109(2). 137-198.

Aronoff, M. 1976. *Word-formation in generative grammar*. Cambridge (MA): MIT Press.

Aronoff, M. & Anshen, F. 2001. *Morphology and the lexicon: Lexicalization and productivity*. In Spencer, A. & Zwicky, A. M. (eds.), *The Oxford handbook of morphology*, 237-247. Oxford: Blackwell.

Audring, J. & Masini, F. 2019. Introduction: Theory and theories in morphology. In Audring, J. & Masini, F. (eds.), *The Oxford handbook of morphological theory*, 1-16. Oxford – New York: Oxford University Press.

Balles I. 2006. *Die altindische Cvi-Konstruktion: Form, Funktion, Ursprung*. Bremen: Hempen.

Balles I. 2008. *The Old Indic* cvi-*construction, the Caland system, and the PIE adjective*. In Rasmussen, J. E. & Olander, Th. (eds.), *Internal reconstruction in Indo-European: Results, methods, and problems*: *Section Papers from the XVI International Conference on Historical Linguistics, University of Copenhagen, 11th-15th August 2003,* 1-15. Copenhagen: Museum Tusculanum.

Bauer L. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer L. 2005. *Theories on productivity*. In Štekauer, P. & Rochelle, L. (eds.), *Handbook of word-formation*, 315-334. Dordrecht: Springer.

Beck, D. 2023. Adjectives. In van Lier, E. (ed.), *The Oxford handbook of word classes*, 365-382. Oxford – New York: Oxford University Press.

Bell, M. J. & Schäfer, M. 2016. Modelling semantic transparency. *Morphology* 26. 157-199.

Berman, H. 1972. *The stem formation of Hittite nouns and adjectives*. Chicago: University of Chicago. (Doctoral dissertation.)

Bertram, R. & Schreuder, R. & Baayen, R. H. 2000. The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy and productivity. *Journal of Experimental Psychology: Learning, Memory and Cognition* 26(2). 489-511.

Bichlmeier, H. 2015. *Caland system and Greek*. In Giannakis, G. (ed.), *Encyclopedia of Ancient Greek language and linguistics*, 256-260. Leiden – Boston: Brill.

Blevins, J. 2016. *Word and paradigm morphology*. Oxford: Oxford University Press.

Bozzone, C. 2016. The origin of the Caland system and the typology of adjectives. *Indo-European Linguistics* 4. 15-52.

Bruening, B. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language* 94(1). 1-42.

Burani, C. & Thornton, A. M. 2003. The interplay of root, suffix and whole-word frequency in processing derived words. In Baayen, H. & Schreuder, R. (eds.), *Morphological structure in language processing*, 157-208. Berlin: De Gruyter.

Bybee, J. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam – Philadelphia: Benjamins.

Bybee, J. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5). 425-455.

CHD = Güterbock, H. G. & Hoffner, H. A. Jr. & van den Hout, Th. P. J. (eds.). 1980. *The Hittite dictionary of the Oriental Institute of the University of Chicago*. Chicago: The Oriental Institute of the University of Chicago.

Comrie, B. 1997. *La famiglia linguistica indoeuropea: Prospettive genetiche e tipologiche*. In Giacalone Ramat, A. & Ramat, P. (eds.), *Le lingue indoeuropee*, 95-122. Bologna: Il Mulino.

Cotticelli-Kurras, P. 1991. *Das hethitische Verbum 'sein'. Syntaktische Untersuchungen* (Texte der Hethiter 18). Heidelberg: Winter.

Cotticelli-Kurras, P. 1998. *Eine lexikalische Untersuchung, Annelies Kammenhuber gewidmet*. In *III. Uluslararası Hititoloji Kongresi Bildirileri: Acts of the IIIrd International Congress of Hittitology, Çorum 16-22 September 1996*, 109-118. Ankara: Uyum Ajans.

Dardano, P. 2007. *In margine al sistema di Caland: Su alcuni aggettivi primari in* *-.nt- *dell'anatolico*. In Groddek D. & Zorman, M. (eds.), *Tabularia Hethaeorum: Hethitologische Beiträge Silvin Košak zum 65. Geburtstag*, 221-246. Wiesbaden: Harrassowitz.

Dardano, P. 2010. *Zur anatolischen Morphosyntax: Das Suffix* -(a)nt- *und seine Bildungen.* In Hazirlaya, Y. & Süel, A. (eds.), *Acts of the VIIth International Congress of Hittitology, Çorum, August 25-31, 2008*, 73-188. Ankara: Çorum Valiliği.

Dardano, P. 2011. *Mehrsprachigkeit im hethitischen Reich: Die Beziehungen zwischen dem Hethitischen und dem Akkadischen*. In Boschung, D. & Riehl, C. M. (eds.), *Historische Mehrsprachigkeit: Workshop des Zentrums für Antike Kulturen des Mittelmeerraums und des Zentrums Sprachenvielfalt und Mehrsprachigkeit an der Universität Köln, 4.-5. Juli 2008*, 47-59. Aachen: Shaker Verlag.

Dell'Oro, F. 2015. *Leggi, sistemi e leghe suffissali "di Caland": Storia della questione "Caland" come problema teorico della linguistica indoeuropea*. Innsbruck: Institut für Sprachen und Literaturen der Universität.

Dixon, R. M. W. 2004. *Adjective classes in typological perspective*. In Dixon, R. M. W. & Aikhenvald, A. Y. (eds.), *Adjective classes: A cross-linguistic typology*, 2-49. Oxford: Oxford University Press.

Dressler, W. U. 2000. *Naturalness*. In Boji, G. & Lehmann, Ch. & Mugdan, J. (eds.), *Morphology* I, 288-296. Berlin – New York: de Gruyter.

Dressler W. U. & Mayerthaler, W. & Panagl, O. & Wurzel, W. U. (eds.). 1987. *Leitmotifs in natural morphology*. Amsterdam – Philadelphia: John Benjamins.

EDHIL = Kloekhorst, A. 2008. *Etymological dictionary of the Hittite inherited lexicon*. Leiden – Boston: Brill.

EHS = Kronasser, H. 1966. *Etymologie der hethitischen Sprache*. Wiesbaden: Harrassowitz.

Francia, R. 2001. La posizione degli aggettivi qualificativi nella frase ittita. In Carruba, O. & Meid, W. (eds.), *Anatolisch und Indogermanisch: Akten des Kolloquiums der Indogermanischen Gesellschaft Pavia, 22./25. September 1998*, 81-91. Innsbruck: Institut für Sprachwissenschaft.

Francia, R. 2020. Hitt. parḫuena-: "oats"?" *Oriens Antiquus* 2. 131-139.

Francia, R. & Pisaniello, V. 2019. *La lingua degli Ittiti*. Milano: Hoepli.

Frotscher, M. 2013. *Das hethitische* -ant- *Partizip und seine indogermanischen Grundlagen: Semantik, Morphologie, Syntax*. Versona: Università degli Studi di Verona. (Doctoral dissertation.)

Frotscher, M. 2017. Review of Maier, M. G. 2013. *Hethitisch* °want- *und Verwan(d)tes: Untersuchungen zur Vertretung des indogermanischen Possessivsuffixes* *-went- *in den anatolischen Sprachen* (Dresdner Beiträge zur Hethitologie 42), Wiesbaden, Harrassowitz. *Wiener Zeitschrift für die Kunde des Morgenlandes* 107. 387-392.

Goedegebuure, P. 2018. *The packagers* -ant- *and* -a-*, and the origin of split-ergativity in Hittite (and Lycian)*. In Goldstein, D. M. & Jamison, S. W. & Vine, B. (eds.), *Proceedings of the 29th Annual UCLA Indo-European Conference*, 77-115. Bremen, Hempen.

Grestenberger, L. 2013. The Indo-Iranian cákri-type. *Journal of the American Oriental Society* 133(2). 269-93.

Grestenberger, L. 2014. *Zur Funktion des Nominalsuffixes* *-i- *im Vedischen und Urindogermanischen*. In Oettinger, N. & Steer, T. (eds.), *Das Nomen im Indogermanischen: Morphologie, Substantiv versus Adjektiv, Kollektivum: Akten der Abeitstagung der Indogermanischen Gesellschaft vom 14. bis 16. September 2011 in Erlangen*, 88-102. Wiesbaden: Reichert.

Grestenberger, L. 2017. *On* *-i- *substantivization in Vedic compounds*. In Hansen, B. S. S. & Hyllested, A. & Jørgensen, A. R. & Kroonen, G. & Larsson, J. H. & Whitehead, B. N. & Olander, T. & Søborg, T. M. (eds.), *Usque ad radices: Indo-European studies in honour of Birgit Anette Olsen*, 193-206. Copenhagen: Museum Tusculanum Press.

Gusmani, R. 1968. *Il Lessico ittito*. Napoli: Libreria Scientifica.

Gusmani, R. 1978. Ittito ḫappina- e greco ΑΠΤΩ. *Incontri Linguistici* 4(2). 242-243.

Haspelmath, M. & Sims, A. 2010. *Understanding morphology*. 2nd edn. London: Hodder Education.

HED = Puhvel, J. 1984-. *Hittite etymological dictionary*. Berlin – New York: de Gruyter.

HEG = Tischler, J. 1983-. *Hethitisches Etymologisches Glossar*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Hill, E. 2020. Analogy in inflectional change: Modification or whole-word replacement? *Language* 96(1). e38-e58.

Höfler, S. 2022. Ein Potpurri aus 'süß' und 'schleimig': "Calandwurzeln", Wortbildung, Etymologie und Co. In Malzahn, M. & Fellner, H. A. & Illés, T.-S. (eds.), *Zurück zur Wurzel: Struktur, Funktion und Semantik der Wurzel im Indogermanischen: Akten der 15. Fachtagung der Indogermanischen Gesellschaft von 13. bis 16. September 2016 im Wien*, 71-82. Wiesbaden: Reichert.

Hoffner, H. A. Jr. & Melchert, C. H. 2008. *A grammar of the Hittite language*. Winona Lake: Eisenbrauns.

Hoffner, H. A. Jr. & Melchert C. H. 2024. *A grammar of the Hittite language*. 2nd edn. University Park, Pennsylvania: Eisenbrauns.

Hrozný, B. 1917. *Die Sprache der Hethiter*. Wiesbaden: Harrassowitz.

HW^2 = Friedrich J. & Kammenhuber, A. & Hoffmann, I. 1975-. *Hethitisches Wörterbuch*. Heidelberg: Winter.

Jie, J. 1994. *A complete retrograde glossary of the Hittite language*. Istanbul: Nederlands Historisch-Archaeologisch Instituut.

Juret, A. 1941. Esquisse d'une vocabulaire étymologique de la langue hittite. *Revue Hittite et Asianique* 6(42). 33-66.

KBo = *Keilschrifttexte aus Boghazköy*, 1916-. Leipzig – Berlin: Hinrichs-Gebr. Mann Verlag.

Kellogg, R. J. 1925. *Studies in Hittite and Indo-European philology.* I. *Some new Indo-European coincidences in Hittite*: Ottawa (KS): Ottawa University.

Kilani-Schoch, M. 1988. *Introduction à la morphologie naturelle*. Bern: Peter Lang.

Kimball, S. E. 2007. *Hittite* ḫūmant- *'all, entire, each'*. In Nussbaum, A. (ed.), *Verba Docenti: Studies in historical and Indo-European linguistics presented to Jay H. Jasanoff by students, colleagues, and friends*, 201-212. Ann Arbor: Beech Stave.

Kloekhorst, A. 2014. *Accent in Hitttite: A study in plene spelling, consonant gradation, clitics, and metrics* (Studien zu den Boğazköy-Texten 56). Wiesbaden: Harrassowitz.

KUB = *Keilschrifturkunden aus Boghazköy*. 1921-1990. Berlin: Akademie Verlag.

Laroche, E. 1963. Comparaison du louvite et du lycien. *Bulletin de la Société de Linguistique de Paris* 58. 58-80.

Lipka, L. 1994. *Lexicalization and institutionalization*. In Asher, R. E. (ed.), *The encyclopedia of language and linguistics*, vol. 4, 2164-2167. Oxford: Pergamon.

LIV^2 = Rix, H. & Kümmel, M. & Zehnder T. & Lipp, R. & Schirmer, B. (eds.) 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihren Primärstammbildungen*. 2nd edn. Wiesbaden: Reichert.

Lowe, J. J. 2014. Indo-European Caland adjectives in **-nt-* and participles in Sanskrit. *Historische Sprachforschung* 127(1). 166-195.

Lundquist, J. & Yates, A. D. 2018. *The morphology of Proto-Indo-European*. In Klein, J. & Joseph, B. & Fritz, M. (eds.), *Handbook of comparative and historical Indo-European linguistics*, 2079-2195. Berlin – Boston: de Gruyter.

Maier, M. G. 2013. *Hethitisch* °want- *und Verwan(d)tes: Untersuchungen zur Vertretung des indogermanischen Possessivsuffixes* *-went- *in den anatolischen Sprachen* (Dresdner Beiträge zur Hethitologie 42). Wiesbaden: Harrassowitz.

Matthews, P. H. 1991. *Morphology*. 2nd edn. Cambridge: Cambridge University Press.

Mayerthaler, W. 1980. *Morphologische Natürlichkeit*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion.

Mayerthaler, W. 1987. *System-independent morphological naturalness*. In Dressler, W. U. & Mayerthaler, W. & Panagl, O. & Wurzel, W. U. (eds.), *Leitmotifs in natural morphology*, 25-58. Amsterdam – Philadelphia: John Benjamins.

Meier-Brügger, M. 2002. *Indogermanische Sprachwissenschaft*. 8th edn. Berlin – New York: de Gruyter.

Meissner, T. 1998. *Das "Calandsche Gesetz" und das Griechische – nach 100 Jahren*. In Meid, W. (ed.), *Sprache und Kultur der Indogermanen: Akten der X. Fachtagung der Indogermanischen Gesellschaft Innsbruck, 22.--28. September 1996*, 237-254. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Meissner, T. 2006. *S-stem nouns and adjectives in Greek: A diachronic study in word-formation*. Oxford: Oxford University Press.

Melchert, C. H. 1984. Review of J. J. S. Weitenberg, *Die hethitischen U-Stämme,* Amsterdam, 1984. *Kratylos* 29. 79-82.

Melchert, C. H. 1994. *Anatolian historical phonology*. Amsterdam – Atlanta: Rodopi.

Melchert, C. H. 2002. Covert possessive compounds in Hittite and Luvian. In Cavoto, F. (ed.) *The linguist's linguist: A collection of papers in honour of Alexis Manaster Ramer*, 297-302. München: Lincom.

Melchert, C. H. 2005. *The problem of Luvian influence on Hittite*. In Meiser, G. & Hackstein, O. (eds.), *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft. 17.-23. September 2000, Halle an der Saale*, 445-460. Wiesbaden: Harrassowitz.

Melchert, C. H. 2017a. *The source(s) of Indo-European participles in *-e/ont-*. In Le Feuvre, C. & Petit, D. & Pinault G.-J. (eds.), *Verbal adjectives and participles in Indo-European languages*, 217-220. Bremen: Hempen.

Melchert, C. H. 2017b. Review of M. G. Maier, *Hethitisch* °want- *und Verwan(d)tes: Untersuchungen zur Vertretung des indogermanischen Possessivsuffixes* *-went- *in den anatolischen Sprachen* (Dresdner Beiträge zur Hethitologie 42), Wiesbaden, Harrassowitz. *Journal of the American Oriental Society* 137(1). 178-179.

Neu, E. 1968. *Interpretation der hethitischen mediopassiven Verbalformen* (Studien zu den Boğazköy-Texten 5). Wiesbaden: Harrassowitz.

Neu, E. 1970. *Ein althethitisches Gewitterritual* (Studien zu den Boğazköy-Texten 12). Wiesbaden: Harrassowitz.

Neu, E. 1974. *Der Anitta Text* (Studien zu den Boğazköy-Texten 18). Wiesbaden: Harrassowitz.

Neu, E. 1979. *Hethitisch* kurur *und* taksul *in syntaktischer Sicht*. In Carruba, O. (ed.), *Studia Mediterranea Piero Meriggi dicata I*, 407-427. Pavia: Aurora.

Neumann, G. 1962. Bemerkungen zur hethitischen Morphologie: Die Verba auf *-eš-*. *Innsbrucker Beiträge zur Kulturwissenschaft* 15. 153-157.

NIL = Wodtko, D. S. & Irslinger, B. & Schneider, C. (eds.). 2008. *Nomina im Indogermanischen Lexikon*: Heidelberg: Winter.

Nussbaum, A. J. 1976. *Caland's "law" and the Caland system*. Harvard University. (Doctoral dissertation.)

Nussbaum, A. J. 2014. *Feminine, abstract, collective, neuter plural: Some remarks on each* (expanded handout), In Neri S. & Schuhmann R. (eds.), *Studies on the collective and feminine in Indo-European from a diachronic and typological perspective*, 273-306. Leiden – Boston: Brill.

Oettinger, N. 1979. *Die Stammbildung des hethitischen Verbums*. Nürnberg: Carl.

Oettinger, N. 1981. Hethitisch *ganenant-* 'gebeugt, durstig': Ein Beitrag zur Heteroklisie. *Münchener Studien zur Sprachwissenschaft* 40. 43-153.

Oettinger, N. 1987. Bemerkungen zur anatolischen i-Motion und Genusfrage. *Zeitschrift für vergleichende Sprachforschung* 100(1). 35-43.

Oettinger, N. 1988. *Hethitisch* -want-. In Neu, E. & Rüster, C. (eds.), *Documentum Asiae Minoris Antiquae: Festschrift für Einrich Otten zum 75. Geburtstag*, 273-286. Wiesbaden: Harrassowitz.

Oettinger, N. 2017. *Der Ursprung des indogermanisches Caland-System und die *-nt-Stämme*. In Le Feuvre, C. & Petit D. & Pinault, G.-J. (eds.), *Verbal adjectives and participles in Indo-European languages: Proceedings of the Conference of the Society for Indo-European Studies (Indogermanische Gesellschaft), Paris, 24th to 26th September 2014*, 267-275. Bremen: Hempen.

Oettinger, N. 2022. *Die Herkunft von indogermaisch* "-u̯ent-". In Malzahn, M. & Fellner, H. A. & Illés, T.-S. (eds.), *Zurück zur Wurzel: Struktur, Funktion und Semantik der Wurzel im Indogermanischen. Akten der 15. Fachtagung der Indogermanischen Gesellschaft von 13. bis 16. September 2016 im Wien*, 225-234. Wiesbaden: Reichert.

Otten, H. & Souček, V. 1969. *Ein althethitisches Ritual für das Königspaar* (Studien zu den Boğazköy-Texten 8). Wiesbaden: Harrassowitz.

Panagl, O. 1982, Produktivität der Wortbildung von Corpussprachen: Möglichkeiten und Grenze der Heuristik, *Folia Linguistica* 16. 225-239.

Pozza, M. 2023. An overview of the Hittite suffix *-ena-/-ina-*: A preliminary assessment. *Rivista degli Studi Orientali* 96(2-3-4). 227-242.

Pozza, M. & Fagiolo, V. 2023. Fragments of scribal bilingualism in lexical lists: The choice of verbal nouns in *-war (-mar)* versus nouns in *-(ā)tar* and *-eššar*. *News from the lands of the Hittites – Scientific Journal for Anatolian Research* 7. 173-193. (*Proceedings of the conference Languages and Cultures in Contact in the Ancient Mediterranean, Verona 11-12 May 2023*, ed. Giusfredi, F. & Martínez Rodríguez, E. & Matessi, A. & Merlin, S. & Pisaniello, V.).

Puhvel, J. 2009. Covert Indo-European verbal roots in Anatolian, *AJNES Aramazd Armenian Journal of Near-Eastern Studies* 4(2). 77-84.

Puhvel, J. 2011. Elliptic genitives and hypostatic nouns in Hittite. *Aramazd Armenian Journal of Near-Eastern Studies* 6(2). 68-72.

Rau, J. 2009. *Indo-European nominal morphology: The decads and the Caland system*. Innsbruck: Institut für Sprachen und Literaturen der Universität.

Rieken, E. 1994. Der Wechsel *-a-/-i-* in der Stammbildung des hethitischen Nomens. *Historische Sprachforschung* 107. 42-53.

Rieken, E. 1999. *Untersuchungen zur nominalen Stammbildung des Hetitische*n (Studien zu den Boğazköy-Texten 44). Wiesbaden: Harrassowitz.

Rieken, E. 2005. Neues zum Ursprung der anatolischen i-Mutation. *Historische Sprachforschung* 118. 48-74.

Rieken, E. 2017. *Das hethitische Partizip: Zur Schnittstelle von Syntax und Semantics*. In Le Feuvre, C. & Petit, D. & Pinault, G.-J. (eds.), *Verbal adjectives and participles in Indo-European languages*, 391-403. Bremen: Hempen.

Rieken, E. & Sasseville, D. 2014. *Social status as a semantic category in Anatolian: The case of PIE* *-wo-. In Melchert, C. H. & Rieken, E. & Steer, T. (eds.), *Munus amicitiae Norbert Oettinger a collegis et amicis dicatum*, 302-314. Ann Arbor: Beech Stave.

Risch, E. 1974. *Die Wortbildung der homerischen Sprache*. 2nd edn. Berlin: de Gruyter.

Sandell, R. P. 2015. *Productivity in historical linguistics: Computational perspectives on word-formation in Ancient Greek and Sanskrit*. University of California. (Doctoral dissertation.)

Starke, F. 1990. *Untersuchung zur Stambildung des keilschrift-luwischen Nomens* (Studien zu den Boğazköy--Texten 31). Wiesbaden: Harrassowitz.

Stüber, K. 2002. *Die primären* s-*Stämme des Indogermanischen*. Wiesbaden: Reichert.

Szemerényi, O. 1985. *Introduzione alla linguistica indoeuropea*. Milano: Unicopli. (Translation of *Einführung in die vergleichende Sprachwissenschft*, 1970, Darmstadt: Wissenschaftliche Buchgesellschaft.)

Tronci, L. 2000. Eredità indoeuropea e innovazione nel greco omerico: L'elemento -ι° come «marca» caratterizzante di primi membri di composto. *Studi e Saggi Linguistici* 38. 275-311.

Watkins, C. 1975. *Die Vertretung der Laryngale in gewissen morphologischen Kategorien in den indogermanischen Sprachen Anatoliens*. In Rix, H. (ed.), *Flexion und Wortbildung: Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, 9.-14. September 1973*, 358-378. Wiesbaden: Reichert.

Watkins, C. 1979. NAM.RA GUD UDU *in Hittite: Indo-European poetic language and the folk taxonomy of wealth*. In Neu, E. & Meid, W. (eds.), *Hethitisch und Indogermanisch: Vergleichende Studien zur historischen Grammatik und zur dialektographischen Stellung der indogermanischen Sprachgruppe Altkleinasiens*, 269-287. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Weitenberg, J. S. 1984. *Die hethitischen U-Stämme*. Amsterdam: Rodopi.

Yakubovich, I. 2006. The free-standing genitive and hypostasis in Hittite. *Journal of Near Eastern Studies* 65. 39-50.

DOI: 10.14746/linpo.2025.67.2.4

A 'norm' in the Ancient Greek dialectal lexicon? A comparative study of Gregory of Corinth's Περὶ Αἰολίδος and the anonymous Γλῶσσαι κατὰ πόλεις (Urb. Gr. 157)*

Wojciech Sowa

Jagiellonian University in Kraków
wojciech.sowa@uj.edu.pl | ORCID: 0000-0002-3425-8435

Abstract: This paper examines how ancient and Byzantine scholars may have conceptualised a "dialectal" lexicon of Greek, with particular attention to the problem of linguistic normativity. It offers a comparative discussion of two markedly different sources: Gregory of Corinth's Περὶ Αἰολίδος and the anonymous lexicographical compilation Γλῶσσαι κατὰ πόλεις (Urb. Gr. 157). Both works seek to illustrate dialectal variation primarily through lexical material, yet they do so within distinct scholarly frameworks and with differing assumptions about linguistic correctness. The analysis draws on criteria developed in modern Ancient Greek dialectology (notably by García Ramón) in order to assess how far the lexical items presented in these sources can be regarded as genuinely dialect specific. Given the secondary nature of the evidence, these criteria cannot be applied mechanically; rather, they serve as a heuristic tool for evaluating the internal logic and reliability of the lexicographical traditions under consideration. Particular attention is paid to the role of literary language, poetic diction, and interdialectal influence in shaping what ancient scholars classified as "dialectal". The study shows that Gregory of Corinth operates with an implicit normative baseline, ultimately rooted in Attic and the learned tradition, against which other dialects are evaluated, whereas the Γλῶσσαι κατὰ πόλεις lack any explicit reference to a standard variety and instead reflect classificatory practices derived largely from literary authority. In both cases, dialectal normativity emerges as prescriptive and scholarly rather than descriptive of vernacular usage. The findings underline the difficulty of defining a "dialectal" lexicon for Ancient Greek and suggest that modern lexicographical approaches must take greater account of the literary, chronological, and scholarly filters through which dialectal material has been transmitted.

Keywords: Ancient Greek dialectology, lexical variation, linguistic normativity, secondary linguistic tradition, literary dialect

* The following paper has been prepared as a part of a research project BPN/BEK/2021/1/00104/DEC/1 financed by NAWA (Narodowa Agencja Wymiany Akademickiej) during a research stay at Departamento de Filología Clásica e Indoeuropeo, Universidad de Salamanca.

LP LXVII (2), 2025. © The Author(s). This is an open access article licensed under the CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en/).

One of the central difficulties in researching Ancient Greek dialectal varieties lies in determining how to interpret the lexical items attested in, or attributed to, specific dialects. In particular, the question arises whether these should be regarded merely as 'Greek words of one region' or whether they ought to be considered 'specifically dialectal' (cf. García Ramón 1999: 524). The criteria proposed – such as those put forward by García Ramón, which attempt to compare attested forms in one dialect with their synonyms in others – may be effective in epigraphical contexts. Such contexts allow for variation in chronology, potential influences from outside the region, and factors such as stylisation mimicking other dialects or poetic idioms. However, these criteria are not operational when examining the more elaborate literary varieties arising in different dialectal regions or the grammatical and lexicographical sources that preserve many rare words (γλῶσσαι) or notable terms (λέξεις). In many cases, these sources transmit words without contextual commentary – for example, without identifying their ultimate source or commenting on their regional distribution. There is also a persistent tendency in scholarship to treat these secondary sources as mere ancillary tools rather than as scholarly works in their own right. Generally, little attention is paid to the reasons why certain lemmas were quoted or included in a lexicon, or to the methodology and linguistic reasoning behind specific *interpretamenta* (Tribulato 2019).

The conception of what constitutes a dialect is clearly crucial in this context. In the case of Ancient Greece, defining the notion of dialect with any precision is notoriously problematic. The existence of different dialects in the first millennium BCE cannot be understood against the background of any form of standard language prior to the emergence of the κοινή. This situation contrasts with that of most modern languages, where one variety frequently attains dominance.[1] Greek dialects are not simply regional variations of a single language, and they differ notably in the chronology of their formation. Nevertheless, "Greek dialects often follow parallel lines of development, but they follow them at different times and different speeds. No correct exegesis of epigraphical – and one may add, literary – texts can be attempted if this is not kept in mind" (Morpurgo Davies 1968: 85).

It also seems likely that the Greeks experienced some degree of 'ethnic unity' despite the evident differences in their speech. They probably possessed an abstract notion of the 'Greek language' (at least by the fifth century BCE; cf. Thucydides, who uses the verb ἑλληνίζειν in the sense of 'to speak Greek') and, as Morpurgo Davies emphasises, "an extensive passive knowledge of different dialects" (Morpurgo Davies 1987: 13). They were certainly aware of

[1] 'dialect: variety of speech differing from the standard or literary language' (OED 599). Compare this with the views of ancient scholars, who recognised only four dialects, aligned with the historical and cultural divisions of the Greek tribes. For example, Clem. Alex. Strom. 1.21.142 states: διάλεκτος δὲ ἐστι λέξις ἴδιον χαρακτῆρα τόπου ἐμφαίνουσα, ἢ λέξις ἴδιον ἢ κοινὸν (4) ἔθνους ἐπιφαίνουσα χαρακτῆρα. φασὶ δὲ οἱ Ἕλληνες διαλέκτους εἶναι τὰς παρὰ σφίσι ε΄, Ἀτθίδα, Ἰάδα, Δωρίδα, Αἰολίδα καὶ πέμπτην τὴν κοινήν, ἀπεριλήπτους δὲ οὔσας τὰς βαρβάρων φωνὰς μηδὲ διαλέκτους, ἀλλὰ γλώσσας λέγεσθαι… 'A dialect is speech that displays the distinctive character of a place, or speech that displays the distinctive – or common – character of a people. The Greeks say that there are five dialects among them: Attic, Ionic, Doric, Aeolic, and, as a fifth, the Koine. The speech of the 'barbarians', however, being unlimited in variety, is not called dialects but languages…' Also note the difference between the terms 'dialect' and 'vernacular'/'accent' (e.g. German *Dialekt* vs. *Mundart*), which in many languages reflects the number of shared isoglosses among smaller dialectal units, contrasted with larger regional groupings.

the existence of distinct regional varieties and displayed various attitudes towards them.[2] The exact situation of spoken, everyday Greek in the Archaic and Classical periods, however, is difficult to reconstruct, as literary sources very rarely preserve genuine dialectal forms or passages. It is nevertheless likely that, in official-political discourse, representatives of individual poleis addressed one another in their own dialects, and that official documents were often prepared in a 'bilingual' (or rather 'bi-dialectal') manner.[3] A separate issue is the use of dialectal features or forms in poetic performance, such as in Attic comedies. One must assume a certain degree of recognition of particular linguistic features as dialectal, which could then be exploited to provoke a comic effect (cf. Colvin 1999 on the use of 'dialects' in Greek literature).

The Attic dialect was certainly not the medium of communication outside the Ionic part of the Greek world before the age of Alexander, even though it exercised a hegemonic role in certain political environments at particular points in time (Bonner 1909: 363). Whether this process can be associated with the concept of 'normativity' – understood as the set of rules, conventions and standards governing language usage within a specific community or region – remains unclear, at least before the rise of grammatical scholarship several centuries later.

If one assumes that linguistic normativity establishes a framework for what is considered correct or acceptable within a given social or cultural context – embracing both formal and informal norms and reflecting conventions developed within a community over time (e.g., grammatical rules or conventions of pronunciation, accent, intonation, vocabulary, idiomatic expressions and politeness strategies) – then it is doubtful whether we may speak of any such framework in Classical times. However, if one instead assumes that the notion of a 'norm' provides a shared set of guidelines, expectations and concepts enabling speakers to communicate effectively and to be understood by others, then the situation appears different.

This paper addresses the broader question of what was – and what should be – considered a genuinely 'dialectal' lexicon, especially in the case of secondary evidence (glosses). It does so through a comparison of two markedly different works: the systematic treatment of the Aeolic dialect by Gregory of Corinth, and the anonymous list of lexical items attributed to various Greek dialects known as the Γλῶσσαι κατὰ πόλεις. Both works attempt to illustrate

[2] Cf. the persistent belief that Greek tribes in remote, less urbanised areas developed distinctive linguistic traits that made their speech difficult for neighbouring groups to understand. For instance, the Eurytanians of mountainous Aetolia were reportedly 'most ignorant of the language and savage in habits' (ἀγνωστότατοι δὲ γλῶσσαν καὶ ὠμοφάγοι εἰσίν; Thuc. 3.94). Their archaic way of life led contemporaries to question – and even contest – their Greek identity.

[3] E.g. Demosthenes notes the difficulty Arcadians and Laconians had in speaking Attic (Dem. Μεγαλοπ. 16.2). Similarly, the vernacular of Lesbos – though celebrated in Aeolic poetry, which was edited, read, and transmitted in Athens during the 5th and 4th centuries BCE – was considered inferior, even 'barbaric' (ἐν φωνῇ βαρβάρῳ τεθραμμένος; Plato Protagoras 341C). For the reception of Sappho, see Yatromanalakis (2007), summarised in Coo (2021: 264-276). While some Athenians may have possessed passive comprehension of other dialects, the extent to which Greeks understood regional varieties remains unclear. Literary sources do not attest individuals fluent in multiple Greek dialects, although multilingualism in foreign languages certainly existed (Bonner 1909: 356-363).

dialectal differences extensively through lexical examples. Whether they do so successfully is, of course, another matter.

In this contribution, I apply the criteria for examining Ancient Greek vocabulary formulated by García Ramón (1999: 522f.; reiterated with slight refinement in García Ramón 2018: 58-60). Particular attention is given to forms that are attested exclusively in one dialect without corresponding forms elsewhere, or that may be specific to one dialect with different, synonymous forms being employed contemporaneously in others. The use of these theoretical premises is challenging, given that the material is not drawn from primary sources. The dialectological analysis serves to assess the general reliability of the sources, and the overarching aim is to clarify expectations concerning what may legitimately be termed a 'dialectal' word.

Gregory of Corinth, Περὶ Αἰολίδος

Gregorios Pardos (Gregory of Corinth) is generally regarded as a representative Byzantine scholar, active in both theological composition and grammatical scholarship, and likely working in the eleventh or twelfth century, although his chronology has long been the subject of debate (cf. Laurent 1963, Becares 1988, Montana 1995: xlviii–xlix, Dickey 2007: 82f.). In addition to his treatises on syntax and rhetorical figures (drawing upon Trypho), and his didactic commentaries on classical literature, he is most widely known for his treatise on the Ancient Greek dialects. This work is the only surviving dialectological treatise transmitted neither anonymously nor with uncertain attribution.[4]

Drawing upon earlier grammatical treatises and traditions (cf. Bolognesi 1953: 101), as well as on his own extensive reading (Bolognesi 1953: 107), Gregory provides an account of the Attic, Ionic, Doric, and Aeolic dialects, formulating general grammatical rules, supplying illustrative examples, and – particularly in his treatment of Attic and Ionic – offering substantial quotations from literary authors (Schoubben 2019: 1-2). He discusses both divergences and affinities among the principal dialects of Ancient Greek, occasionally referring to 'minor' varieties such as Boeotian and Thessalian. His work situates itself within the established grammatical tradition, exemplified by figures such as Dionysius Thrax and Apollonius Dyscolus, who classified dialects primarily on phonetic, morphological, and syntactic grounds. Gregory nonetheless introduces several original elements, including the systematic use of examples drawn from both poetry and prose, the comparison of dialectal forms with what he treats as standard Attic, and the explicit evaluation of dialectal usage in terms of correctness and stylistic elegance (cf. Wilson 1996: 188).

In the chapter devoted to Aeolic (specifically the Lesbian variety, *kleinasiatisch-aiolisch*), the dialect is described in a formalised and schematic manner. Gregory presents a series of rules defining the dialect, drawing attention to those features that he interprets as deviations from an underlying linguistic 'norm'. The structure of this account reflects its

[4] Cf. the presentation of the oeuvre in Kominis (1960), Donnet (1966, 1967), Wilson (1983a: 184-90), Bolognesi (1953), Glucker (1970), Montana (1995), Robins (1993: 163-72), and Hunger (1982).

principal source, namely the fragments of Johannes Grammaticus (Philoponus) preserved in the *Compendia Περὶ Αἰολίδος* (for the textual tradition see Hoffmann 1897: 204-222). Priority is accorded to the phonetic and phonological system, followed by the morphology of nouns and verbs. A final section lists lexical items which, in Gregory's view and according to earlier authorities, are peculiar to the Lesbian dialect. These appear in §§57-67, cf.

57	Τὸ ἐγὼ περισπῶσιν, ἐγῶν λέγοντες, καὶ ἐμὼ ἐμών	'The form ἐγὼ is twisted into ἐγῶν, and ἐμὼ becomes ἐμών.'
58	Τὸ σὺ τὺ ὁμοίως τοῖς Δωριεῦσι λέγουσι	'The form σὺ corresponds to τὺ, as the Dorians also say.'
59	Καὶ τὸ ἡμῖν ἄμμιν, ἡμεῖς ἄμμες	'Similarly, ἡμῖν corresponds to ἄμμιν, and ἡμεῖς to ἄμμες.'
60	Τὸ ἐνθάδε, τῇδε	'The form ἐνθάδε corresponds to τῇδε.'
61	Τὸ ἄλλοσε, ἑτέροσε	'The form ἄλλοσε corresponds to ἑτέροσε.'
62	Τὰ χείρονα, χέρρονα	'The form χείρονα appears as χέρρονα.'
63	Τὸ ἐπαινῶ, αἴνημι	'The verb ἐπαινῶ is rendered as αἴνημι.'
64	Τὰ ἱμάτια, ἔμματα	'The word ἱμάτια appears as ἔμματα.'
65	Ἡ μία, ἴα	'The feminine form μία corresponds to ἴα.'
66	Ἡ ψῆφος, ψᾶφαξ	'The noun ψῆφος is rendered as ψᾶφαξ.'
67	Τὸ ἔστρωται, ἐστόρηται	'The verb ἔστρωται appears as ἐστόρηται.'

Although this structure is not markedly different from that employed in many modern treatments of the Greek dialects, it must be emphasised that Gregory's categorical distinctions are frequently imprecise. Lexical items are at times deployed to abstract phonological or morphological rules, and certain purely morphological alternations are misinterpreted as dialectal phonetic processes. Moreover, some lexical items (e.g. §22 ἔδοντας· ὀδόντας, ἐδύνας· ὀδύνας) are cited despite now being recognised as morphological or phonological variants rather than genuine dialectal forms.

A close examination of the lexical section of the Aeolic chapter demonstrates that only four of the ten items cited may be regarded as authentically dialectal. These are primarily associated with the poetic register of the dialect – a predictable outcome, given that the sources upon which Gregory and earlier lexicographers rely are largely literary.

ἄμμες (§ 59)

One such form, frequently cited as characteristic of the Lesbian dialect, is the nominative plural ἄμμες 'we'. Gregory presents it together with the dative plural ἄμμιν (Καὶ τὸ ἡμῖν ἄμμιν, ἡμεῖς ἄμμες), thereby conflating two observations in the *Compendium* of John Philoponus (Comp. III §§60 and 61). This brief entry conveys linguistically reliable information with a solid foundation in both inscriptional and literary Aeolic (Buck 1955: 98ff., Thumb & Scherer 1959: 99ff.; cf. the attestations in Hamm 1957: 107 and Voigt 1971: 385).

In the epigraphical record the first-person plural pronoun appears, for example, in the genitive αμμεων and dative αμμι,[5] and functions also as a possessive adjective, as in the formula τας αμμας πολιος (MAT 010, 35; ERE 010, 73.90; MYT 163, 14), which is attested as late as the Tiberian period.[6]

Both ἄμμες and ἄμμιν occur in Homer[7] and Hesiod,[8] and appear also in non-epic poetic authors such as Pindar (ἄμμιν 4×, e.g. *Hymn* 42.2 μόχθος ἄμμιν – τοῦτό γέ τοι ἐρέω) and Theognis (ἄμμιν 2×, e.g. *Elegiae* 418 χρυσός, ὑπερτερίης δ' ἄμμιν ἔνεστι λόγος). In the ancient grammatical tradition these forms are consistently assigned to Aeolic, e.g. Herodian (Περὶ ὀρθογραφίας 3.2.517.14f.): ἡμεῖς· ἀμές γὰρ λέγουσι οἱ Δωριεῖς καὶ ἄμμες οἱ Αἰολεῖς.[9] The entire Aeolic paradigm (i.e. Lesbian and Boeotian) is discussed in detail by Apollonius Dyscolus (*De pron.* 93.23-103.12). A geminated form of ἄμμες is even erroneously described as Doric in certain scholia (cf. *Schol. Hom. Il.* 8.352.1–3. 1-3 οὐκέτι νῶϊ] κοινῇ ἡμῖν- Δωριεῖς γάρ φασιν ἄμμες, Ἀττικοὶ δὲ νώ, Ἴωνες ἡμέας. ἔστι δὲ ἀντωνυμία δυϊκὴ ὀρθῆς καὶ αἰτιατικῆς πτώσεως. ἐὰν δὲ σὺν τῷ ν, ἡμῶν καὶ ἡμῖν).

The dialectal evidence appears to confirm the use of /a/ and of a geminate /m/ in Lesbian. The personal pronouns of the first- and second-person plural attracted particular interest among ancient grammarians owing to their presence in Homer and in Lesbian lyric. In the Homeric poems two plural paradigms coexist: one based on ἄμμες, conventionally labelled 'Aeolic' (with aspiration and recessive accent), and one based on ἡμεῖς, conventionally labelled 'Ionic'. The Aeolic series comprises nom. ἄμμες, acc. ἄμμε, dat. ἄμμι(ν), with no distinct genitive (only ἡμέων is attested). The Ionic series includes nom. ἡμεῖς, acc. ἡμέας, and dat. ἡμῖν (Chantraine 1948: 268; for additional forms, such as acc. ἧμεας/ἧμας, dat. ἧμιν/ἧμιν, and the artificially extended gen. ἡμείω, cf. Chantraine 1948: 268-271).

Although both Aeolic and Ionic forms derive from the same proto-form, they reflect different outcomes of the so-called 'first compensatory lengthening'. Aeolic diverges markedly from the Ionic-Attic treatment of *-*Vsm*-, yielding a short vowel plus a geminate (Lejeune 1972: 122f., Thumb & Scherer 1959: 95f., Scarborough 2023: 131-134). Thus

[5] Cf. gen. μεταπεμπομενων αμμεων και Ερεσιων δικαστηριον MAT 010.29 (mid-2nd century BC; cf. also MYT 026.8, 11; 209.4; 225.15); παραγεγοναν εις ταμ πολιν αμμων ERE 010.61; dat. εοντες αμμι συγγγε[νεες] LES 01.8 (late 3rd century BCE; cf. also MAT 010.30.51; ERE°010.68.96); [απεστελλεν] αμμι LES 05.5 (1st cent. AD); δεδοσθαι δε (...) πολιτειαι παρ' αμμιν MAT 010.48.

[6] Cf. also [π]ροξενος τας πολιος αμ[μ(ε)ων] EOL 05, c. 3 (late 3rd cent. BCE); κτ<ι>σταν τας πολιος αμμεων MYT 225.15 (1st century AD). See also the catalogue of forms in Hodot (1990: 134). A similar situation occurs in Thessalian and Boeotian: Thess. gen. αμμεουν, acc. αμμε; Boeot. gen. αμεων (αμιων Ap. Dysc. *De pron.* 95.21; Blümel 1982: 267).

[7] ἄμμες occurs 4×, e.g. Il. 21.432 τώ κεν δὴ πάλαι ἄμμες ἐπαυσάμεθα πτολέμοιο; ἄμμιν occurs 4×, e.g. Il. 13.379 Ἄργεος ἐξαγαγόντες ὀπυιέμεν, εἴ κε σὺν ἄμμιν, but ἄμμι appears as many as 16×.

[8] E.g. *Scutum* 88: ᾗ ἀλόχῳ – τάχα δ' ἄμμες ἐπιπλομένων ἐνιαυτῶν, etc.

[9] Also Apoll. Soph. *Lex. homericum* 24.26-28: ἄμμε Αἰολικῶς ἀντὶ τοῦ ἡμᾶς· […] τὸ δὲ ἄμμι ἡμῖν. Eustath. *Comm. ad Hom. Il.* 1.172.26: εἰμί ἐμμί· ἡμεῖς ἄμμες. Eustath. *Comm. ad Hom. Il*. 4.73.16–74.4: τὸ ἡμῖν ἄμμι λέγουσιν οἱ Αἰολεῖς, βαρύνοντες αὐτὸ καὶ συστέλλοντες τὴν λήγουσαν. Ἀπολλώνιος: «ἄμμι γε μὴν νόος ἔνδον ἀτύζεται». Δωριεῖς δὲ ἀμίν, συστέλλοντες τὸ ι καὶ ὀξύνοντες […] ὅτι τὸ ἡμεῖς ἅμες λέγουσιν οἱ Δωριεῖς, ἄμμες δὲ οἱ Αἰολεῖς.

**n̥s-mé* > **asmé* > Aeolic /amme/ ἀμμέ[10] stands in contrast to Ionic-Attic /a:me/ ἄμέ-, whence ἡμέ-, to which case endings were subsequently added: **n̥s-mé-es* > ἡμεῖς in Ionic--Attic; **n̥s-mé-s* > ἄμές (Doric) and ἄμμες (Lesbian) (Rix 1992: 178, Dunkel 2004: 18ff.).

It is plausible that the retention of both sets of forms in the Epic tradition reflects metrical constraints: Lesbian ἄμμε(ς) is metrically equivalent to Proto-Ionic *ἡμέ(ς) (Wyatt 1992: 170f.). Yet there remains a substantial difference between the two series, e.g. ἄμμες vs. ἡμεῖς and ἄμμε vs. ἡμέας. Ionic forms tend to be spondaic, whereas Lesbian forms are predominantly trochaic (particularly in the third foot) and may even be monosyllabic. As Miller observes, dialectal (non-Ionic) forms may in certain cases occupy analogous metrical positions, e.g. ἦμιν (8×) ~ ἄμμι(ν) (7×) at the beginning of the third foot, and ἡμῖν (21×) ~ ἄμμι(ν) (4×) at the beginning of the sixth foot. In the latter case, the Aeolic forms may serve to raise the stylistic level of the verse (Miller 2014: 316; cf. Meier-Brügger 1986: 127-143 for a comprehensive documentation and interpretation of Aeolic pronoun forms in Homer). Yet the question remains whether the 'Aeolic' forms of the personal pronoun in Homer – despite being normalised in accentuation and psilosis under the influence of the transmission of Sappho and Alcaeus – should be interpreted as genuine survivals of an Aeolic linguistic substrate and of an 'Aeolic phase' in the development of Greek epic, or whether they are in fact archaisms inherited from the pre-alphabetic stage of the tradition. A definitive answer remains elusive.[11]

αἴνημι (§ 63)

In paragraph 63 (Τὸ ἐπαινῶ, αἴνημι), Gregory cites the athematic verb form αἴνημι as specifically Aeolic in contrast to the thematic αἰνέω, which is the regular Greek form (cf. e.g. *Hom.* αἰνεῖτ' Θ 9). The verb αἰνέω ('to praise, glorify, approve') is frequent in poetic diction but appears only rarely in Attic prose, where the derivative ἐπαινέω is strongly preferred.

The form αἴνημι does not occur in the surviving Lesbian lyric corpus, nor is it attested in epigraphic material from Lesbos, Thessaly, or Boeotia. From a formal perspective, it represents a shift from the class of so-called *verba vocalia* (contract verbs) to athematic forms – a development characteristically associated with the Aeolic dialect (cf. Scarborough

[10] With /am/ < *[n̥] (Rix 1992: 66).

[11] See Miller (2014: 317), who assumes that a treatment similar to the Aeolic 'first compensatory lengthening' can already be observed in Mycenaean and in archaic Arcadian forms from the Peloponnese. Cf. the document from Megalopolis, 207/6 BCE (I.v. Magn. 38, see Dubois 1986 II: 273 ff.): line 24: εχοντες ευνοως προς αμμε; line 22: προς παντας τος κα αμε εαρημενος; line 18: ται δε πολι ται αμετηραι αμε (accusative 'us'). Another interpretation is offered by Dubois (1986 I: 79), according to whom αμε is a regular form with lengthened /a:/, whereas αμμε represents an Aeolic form erroneously recorded by an Ionic scribe. If in Mycenaean there indeed existed a treatment analogous to the Aeolic (geminated sonant; compare a-ke-ra$_2$-te PY Vn 493.1 /agerrantes/ or /angellantes/, see Peters 1986: 3068, 313), then one may also postulate the existence at this stage of the form ‡ammé. In this scenario, the Homeric ἄμμε could similarly be regarded as an archaic Mycenaean element. Within the Mycenaean-Ionic poetic tradition, ἀμμέ(ς) may have been preserved precisely because of its metrically distinctive structure. Subsequently, forms such as ἀμμέ(ς), later modified to ἄμμε(ς), entered poetic usage under the influence of Lesbian poets. For ordinary speakers, however, the archaic and artificial ἀμμέ(ς) was naturally replaced by forms still current in the living dialect of the Aeolians of Asia Minor.

2023: 136-142). The grammatical tradition itself acknowledges this tendency, citing the pair αἰνέω : αἴνημι alongside forms such as φιλῶ : φίλημι.[12]

The verb αἰνέω functions in parallel with the thematic noun αἶνος 'story, tale' (*Il.* Ψ 652), which could suggest that, as in many other cases of */-eo:/*-forms, we are dealing with a denominative formation (cf. Risch 1974: 300-319). The situation is, however, complicated by the aorist formation: Ionic-Attic αἰνέω constructs an aorist in -ε(σ)σα, whereas Homer consistently employs a form in -ησα.

Morphologically, verbs of the type φιλέω – ήσω – ῆσα derive from *-o*-stems (whence *αἶνος* might be expected), while the aorist in -ε(σ)σα is largely restricted to *-es*-stems (cf. e.g. αἰδέομαι). Consequently, it is difficult to determine decisively whether αἰνέω is a denominative formation or whether αἶνος and related nouns are instead back-formations from the verb (cf. Tucker 1990: 37[25], 94). Matters are further complicated by the fact that αἰνέω itself lacks a convincing etymology (cf. Frisk 1960: 40-41, Beekes 2010: 39-40). Comparative evidence from Tocharian and Hittite may, however, appear promising – cf. Hitt. *enant* 'tame' (MAŠ.GAL *enanza* 'tame goat'), an intransitive participle in *-ant-* from *en-* (< *ain-* 'to be agreeable'; Puhvel 1984 II: 271), and Toch. A/B *en-* 'to instruct, teach, chastise', e.g. *tumeṃ lyama asānne enṣṣate-me* 'then he sat down on his seat and taught them' (Adams 2013: 87). Should these comparanda indeed relate to Greek αἰνέω (αἶνος), they must be explained as continuants of PIE **h₂eǐ-n-* (Peters 1980: 80).

Setting aside these etymological questions, it seems clear that Gregory quotes αἴνημι because of its morphological character, which the ancient tradition viewed as typically Aeolic. Interesting in this regard is its single literary attestation outside Homer, namely *Hes. Op.* 683: οὔ μιν ἔγωγε || αἴνημ', οὐ γὰρ ἐμῷ θυμῷ κεχαρισμένος ἐστίν. This form has been interpreted as a specifically Hesiodic Aeolism (Thumb & Scherer 1959: 8), given that the Iliad and Odyssey consistently employ the thematic forms. It almost certainly entered the grammatical tradition via the scholiastic commentaries; cf. *Schol.* in *Op.* (Prolegom. Schol. 681.1-2): AINHM'. Αἴνημι, αἰνῶ, καὶ κατὰ παραγωγὴν Ἀττικὴν αἴνημι.

ἔμματα (§ 64)

The noun ἔμμα, meaning 'dress, garment', is cited twice by Gregory: in paragraph 64 and again in paragraph 20. The form appears in grammatical and lexicographic sources; cf. *Hsch.* E 2355 ἔμμα· ἱμάτιον 'garment', and *Hsch.* Γ 319 γέμματα· ἱμάτια, which preserves an orthographic tradition employing Γ rather than the expected *F (*ϝέμματα). For the spelling <Γ> in place of digamma <F> and the historical development of PIE */w/ in Aeolic dialects, see Sowa (2011: 166-167). Herodian explicitly classifies the word as Lesbian, cf. *Περὶ ὀρθ.*

[12] E.g. Eustath. *Comm. in Hom. Od.* 1.80.11-13: ὁμοίως τῷ διζῶ δίζημι […] καὶ τῷ φιλῶ φίλημι […] καὶ τῷ αἰνῶ αἴνημι… Od. 2.247.31 ff.:ώς τοῦ αἰτῶ τὸ αἴτημι καὶ τοῦ φιλῶ τὸ φίλημι, καὶ τοῦ αἰνῶ τὸ αἴνημι…: "Just as διζῶ gives rise to δίζημι, and φιλῶ to φίλημι, so αἰνῶ gives rise to αἴνημι; …. likewise, αἰτῶ forms αἴτημι, φιλῶ forms φίλημι, and αἰνῶ forms αἴνημι."

3.2.500.28: εἶμα τὸ ἱμάτιον διὰ τῆς ει διφθόγγου· ἔμμα γάρ φασιν οἱ Αἰολεῖς.[13] Likewise Compendium III §54 lists τὰ ἱμάτια ἔμματα as Aeolic.

Although the word does not occur in surviving Lesbian inscriptions, it appears three times in the transmitted Lesbian poets, e.g. *Alc.* fr. 58.21: [ ]τ' ὀνάρταις χέρρ' ἀπύ μ' ἐμμάτων; *Sapph.* fr. 62.12: τὰ τ' ἔμματα κα[; *Sapph.* fr. 276 (2) col. 3.41: κων.[.]φιλ.[Ἐρίγυιοσ[]πὲρ ἐμμάτ[ω]ν. It is also classified as an Aeolism in *Lyrica Adespota* fr. 9.2 (Ps.-Alcman): παίσαι παρθενικαί, παίσαι καλὰ ἔμματ' ἐχοίσα[ι].

Related forms appear in Doric, particularly in Crete (e.g. *IC* IV Gortyn 72 V 40; 75 B 3: ϝεῖμας, ϝεῖμας κ' ἀνπιδέμας), and in East Ionian inscriptions, e.g. Chios 505 (4th c. BCE): ανηρ εμα πυρι κατεκα, where εμα is plausibly equivalent to εἶμα (Hernández Vázquez 1994: 205).[14]

ἔμμα is a regular equivalent of Ionic εἶμα /heːma/ and should be interpreted as a neuter formation with the suffix *-mn̥* (cf. Risch 1974: 49f., Rix 1992: 33), compare OInd. *vásma*. The derivational base is the verbal root *u̯és-* (cf. Greek ἕννυμι < *u̯es-nu-mi*) 'to be clothed, to dress oneself',[15] also attested in other Indo-European languages, such as Hittite *wēsta-* 'to be dressed in (something)' and Vedic imperfect *váste* 'he had on' (see LIV[2]: 692f. for comparative material).

From the perspective of dialectal phonology, ἔμμα exhibits the typical Aeolic treatment of the consonant cluster */sm/, which regularly develops into the geminate /mm/. The orthography found in lexicographical sources, using gamma <Γ> instead of digamma <ϝ>, can be considered a result of textual transmission. Although *(γ/ϝ)ἔμμα* does not appear in epigraphic material from Lesbos or neighbouring areas of Asia Minor, the form should still be interpreted as part of the actual dialectal layer in Lesbian lyric poetry. It cannot be considered an element of the Homeric epic tradition, as all examples of the noun in epic consistently display the expected Ionic form εἶμα.

The plural ἔμματα is attested in poetry, for example Sappho fr. 62, 12 (τά τ' ἔμματα κα[…]) and in Pseudo-Alkman (*Lyr. Adesp.* fr. 9, 2), in a distinctly dialectal literary verse, yielding traces of the Lesbian variety (παίσαι παρθενικαί, παίσαι καλὰ ἔμματ' ἐχοίσα[ι]), with a dialectal treatment of the group */ns/ – παίσαι (instead of πᾶσαι), ἐχοίσα[ι] (for ἔχουσαι). This evidence further supports the interpretation of ἔμμα/ἔμματα as typical dialectal forms.

[13] Cf. the context 28 ff.: εἶμα τὸ ἱμάτιον διὰ τῆς ει διφθόγγου· ἔμμα γάρ φασιν οἱ Αἰολεῖς ἀποβάλλοντες τὸ ι καὶ διπλασιάζοντες τὸ σύμφωνον, ὥσπερ κείρω κέρρω, φθείρω φθέρρω. τὸ δὲ ἱμάτιον καὶ τὰ παρ' αὐτὸ πάντα διὰ τοῦ ι γράφεται, οἷον ἱματίζω, ἱματισμός, ἱματιοφόρος καὶ ἱματιοπώλης… 'The word εἶμα ('garment') is written with ι rather than ει because of the way Aeolic treated diphthongs. The Aeolians, it is said, would drop the ι and double the following consonant, as in κείρω → κέρρω or φθείρω → φθέρρω. All words related to ἱμάτιον – such as ἱματίζω, ἱματισμός, ἱματιοφόρος, and ἱματιοπώλης – are likewise written with ι.'

[14] Cf. however the use of ἱμάτιον in the East-Ionic dialect (Samos 3x, 346-345 BCE, Samos 133 = IG XII, 6 1:261, lines 27, 31,33) ἱμάτιον λευκόν, ἡ ὄπισθε θεὸς ἔχει, ἱμάτια Ἑρμέω 'A white garment, which the goddess holds behind her; garments of Hermes' (Hernández Vázquez 1994: 338).

[15] This nasal infix present is a Greek innovation, not attested in other IE languages, cf. however, Hom.-Ion. pres. εἶμαι < **u̯és-mai̯*, used as a perfect to ἕννυμι, which should, therefore, be considered more archaic (LIV[2]: 692 f.); cf. also Hsch. A 1363: ἄεμμα· τόξον (Call. *Hymn.* 2,33), ἱμάτιον.

Therefore, it appears that the information provided by Gregory is accurate and that the quoted form is genuinely dialectal. Nevertheless, doubts remain regarding the context and interpretation proposed by the author. From a modern perspective, one cannot claim that a graphic sign such as iota (or its equivalent) changes into another letter, for example to represent a liquid or nasal sound. Rather, the 'rule' offered by the author should be understood as an application of the ancient theory of πάθη (modifications), whereby one letter could be substituted for another to explain a word's etymology. This paragraph, which groups various forms as the result of different phonological laws, is a typical example of such a 'change' (μετάληψις; cf. Dickey 2007: 246, 250).

ἴα (§ 65)

In this paragraph Gregory introduces the Lesbian feminine form of the numeral 'one', ἴα, in place of the Ionic-Attic form μία. This observation is supported in the grammatical tradition and is quoted directly from §55 of *Compendium* III. Gregory correctly situates the form within the Aeolic context, as the feminine ἴα/ἵα constitutes an isogloss uniting Boeotian, Thessalian, and Lesbian (Buck 1955: 94, Thumb & Scherer 1959: 4, Blümel 1982: 271).[16] The psilotic Lesbian variant ἴα is regarded as a secondary innovation, comparable to Boeotian forms such as the genitive ιας (SEG 3:359.10) and the accusative ἴαν in Corinna;[17] and to Thessalian accusative ιαν (IG IX² 517.22, 44; SEG 13:3958). In Lesbian, the form is attested only in literary contexts (acc. sg. ἴαν in Sappho 56.1 and Alcaeus fr. 350.6). The epigraphic record is sparse: Mytilene yields one nominative form, μηδεια (MYT 04, 12), alongside Koine forms [ουδε]μιας (MYT 024.56, 2nd c. BCE) and [μ]ηδεμιας (MYT 026.10, 12, 1st c. CE), which renders its interpretation ambiguous. It is equally plausible that μηδεια represents a feminine form of the adjectival pronoun ἴος (Hodot 1990: 152; literary attestations in Hamm 1957: 109).

Forms of this type also occur in Homeric Epic: nominative ἴα (Δ 437), accusative ἴαν (ξ 435), genitive ἰῆς (O 173, Ω 496), and dative ἰῇ (I 319). A metrically motivated neuter dative ἰῶι (Z 422) is also attested. This raises the question of whether such forms represent dialectal innovations (perhaps indicating an 'Aeolic' element in Epic diction; cf. Thumb & Scherer 1959: 210), archaisms, or analogical developments. According to García Ramón, reflexes of the inherited feminine **smih₂-* (cf. Arm. **mi* < **smi̯a-*) lost the initial */m/ to align the feminine with the masculine and neuter forms εἷς and ἕν (García Ramón 1975: 65; similarly, Ruijgh 1971: 601). Yet this model of proportional analogy is problematic – it presupposes the development of ἵ- /i-/ from masculine nominative εἷ- /he:-/ or oblique ἑ- /he-/, instead of the historically expected †ἕα (< εἷς, †ἕα, ἕν). The motivation for such alignment is unclear (Parker 2008: 448; cf. already Schmidt 1900: 391-399), particularly as other Greek

[16] The consistent placement of ἴα in editions of Boeotian and Thessalian inscriptions is influenced by the Lesbian and Homeric ἴα. This contradicts both the etymology of the form and the phonology of Thessalian and Boeotian, neither of which are psilotic dialects (Scarborough 2023: 122[240]).

[17] Cf. τὰν δ' ἴαν Μή[ας] ἀγαθὸς (Corinna iii.17; cf. Page 1953: 57).

dialects did not attempt to regularise the paradigm and continued to employ the suppletive feminine μία.

An alternative explanation for the emergence of the feminine form ἵα/ἵα has been proposed by Parker, building on Schmidt's assumptions. Parker argues that, due to phonological developments between the Proto-Indo-European (PIE) period and the Proto--Greek stage, the inherited feminine form took the shape nom./acc. **smi̯a-*, obl. **si̯a-*. This allomorphism later gave rise to the forms μία/ἵα through analogical alignments during the formation of the Greek dialects (Parker, ibidem). He proposes a generalisation of the zero-grade form **sm-i̯éh₂* from dependent cases at an early stage. The disappearance of the */m/* segment between the spirant */s/* and the non-vocalic */i̯/* resulted in **si̯éh₂-,* as seen in Hittite *šī-* 'one' (Kloekhorst 2008: 750f.), reflecting a parallel development in the Indo-Iranian pronominal system. In Proto-Greek, this process produced the allomorphism **smi̯a-* in the nominative/accusative and **si̯a-* in the dependent cases. In the nominative/accusative, regular assimilation yielded **sm-* > **m-* (μία, μίαν; see Lejeune 1972: 120f.), whereas the development of ἵα- from **si̯a-* is less straightforward. It is likely that the sequence **si̯-* assimilated to **i̯i̯-*, analogous to the treatment of the intervocalic group **-si̯-*, producing forms such as ἰᾶς, ἰᾷ (or, via metathesis, **i̯i̯a-* > **iha-* > **hia-*).

Alternatively, the Lindeman variant suggests that monosyllabic forms developed into bisyllabic forms (**smi̯a-/*si̯a* > **mii̯a-/hii̯a*), comparable to bisyllabic variants attested in the Balkan area, for example in Albanian and Armenian.[18] Thus, at the stage of dialectal differentiation, Greek inherited an allomorphic alternation **smi̯a-* : **si̯a-*. The Aeolic and Homeric forms ἵα/ἵα are best understood as levellings from dependent-case **si̯a-* to the nominative/accusative, whereas μία, μίαν exhibit the opposite direction of analogical spread. Gregory, therefore, correctly identifies a noteworthy phenomenon, albeit one limited in its distribution. The rarity of ἵα/ἵα in Greek is unsurprising – analogical levelling typically proceeds from the main cases to the dependent ones and not vice versa (Scarborough 2023: 122-126).

The remaining forms adduced by Gregory as exclusive lexical items cannot, for several reasons, be regarded as dialectal. Many belong to the literary dialect, others derive directly from Homeric diction, and several are unattested in epigraphic or literary sources. Their inclusion as 'typical' Lesbian forms is therefore difficult to justify.[19] The examples cited by Gregory reflect a strong dependence on the grammatical and literary tradition, which serves as his principal source. Throughout the treatise, Gregory demonstrates a consistent preference for Attic as the most refined and authoritative dialect of Greek. Attic is implicitly

[18] Arm *mi,* gen. *mioy* < **smii̯o-*; Alb. *një* < PAlb **smi̯o-s* < IE **smii̯o-*, both form understood as secondary masculine formations derived from feminine **sm(i)i̯a* < **smi(i̯)ə-* < **smih₂-*. This has been interpreted as one of the isoglosses supporting the assumption of a Balkan-Indo-European linguistic area (Matzinger 2012: 151).

[19] See, for example, paragraphs 60 and 61, where Gregory draws attention to allegedly dialectal forms of the adverbs τῇδε and ἑτέροσε, which, however, cannot be attested for Lesbian. In both epigraphic and literary evidence, the alternative forms τυῖδε and ἄλλυι/ἄλλοι should be considered the genuinely dialectal ones (Sowa 2024: 229-231). This observation is corroborated by the evidence and observations of ancient grammarians and lexicographers (e.g., Hsch. T 1615 notes: τῦδε- ἐνταῦθα. Αἰολεῖς; *Schol. In Il.* Ξ 298 μήποτε δὲ ἰσοδυναμεῖ τῷ παρ' Αἰολεῦσι 'τυῖδε'; Alc. Fr. 392 οὐδέ τι μυνάμενος ἄλλοι τὸ νόημμα). For the formation of adverbs in the Lesbian dialect, see Bechtel (1921: 103), Buck (1955: 103), Thumb & Scherer (1959: 109), Hamm (1957: 113), Risch (1974: 358), Rix (1992: 170), Rodríguez Somolinos (1998: 101, 205).

treated as the normative standard against which other dialects are evaluated; hence, the frequent remarks such as "contrary to Attic" or "against Attic usage". Gregory occasionally criticises forms diverging from Attic norms ("they… falsely… with regard to pronunciation"). He praises Attic for its richness, clarity, and harmony, attributing its excellence to the cultural achievements of Athens in literature, philosophy, and politics. He regards Attic as the language of educated Greeks and advises his students to imitate its style and vocabulary. The normative role of Classical Attic is thus maintained, even though Gregory is aware that the linguistic standard of his own day is the Byzantine κοινή (cf. remarks such as "in our dialect/use").

However, Gregory does not dismiss the other dialects entirely. He recognises that different dialects are suited to different genres: Ionic for Epic, Doric for choral lyric, and Aeolic for erotic poetry. He allows some degree of contextual variation and acknowledges that dialects possess their own merits and peculiarities. He also accepts that Doric and Aeolic appear to share elements of a common linguistic history distinct from Ionic or Attic and even proposes that the Lesbian dialect is a precursor of Latin (cf. §29 τοῦτοις… καθὰ δὴ καὶ οἱ Ῥωμαῖοι, τούτων ὄντες ἄποικοι). Overall, the material he cites reflects a tendency – characteristic of late antique and Byzantine scholarship – to interpret archaic or irregular forms, especially those found in early poetry, as dialectal, and often specifically as Aeolic.

Γλῶσσαι κατὰ πόλεις

The anonymous lexical list transmitted under the title Ποῖαι γλῶσσαι κατὰ πόλεις constitutes an intriguing document relevant to the study of dialectal vocabulary. It presents a list of one hundred forms assigned to various dialects, regions, and even specific cities, and is preserved in Urb. gr. 157 of the Vatican Library. Bekker published the text in *Anecdota Graeca* III (1095–1096) without commentary, and subsequent scholars of the Greek dialects – notably Hoffmann, Meister and Bechtel – have consulted the list principally as supplementary material for regional vocabularies. A critical edition is still lacking (cf. Latte 1925: 136 for the essential information on its textual tradition). Beyond the Vatican witness, at least two further copies survive in Madrid (Royal Library, cod. XL and cod. XCV; cf. Iriarte 1769, 146 and 378), together with three later exemplars in Italian libraries.[20] With the exception of short studies by Latte (1925: 136-147) and Bowra (1959: 43-60), the document has attracted little scholarly attention; the occasional references that do occur typically treat the list as a reliable source of dialectal evidence (e.g., Peters 1994: 210).

[20] The two copies preserved in Madrid were transcribed within the circles of disciples of the Greek humanist Konstantinos Laskaris and form part of codices containing various grammatical writings, including Philoponus and the grammar prepared by Laskaris himself. Two further manuscripts are held in Italian libraries: Bibl. Laur. Plut. 58.19, fol. 188r-191 (copied by Francesco Filelfo), and Venice, Biblioteca Nazionale Marciana, gr. X. 026 (coll. 1305), fol. 202v-203v. An additional copy of the list by Filelfo is preserved in the Biblioteca Ambrosiana in Milan (Ambr. F 14 sup., ff. 161v-165v) – an exemplar unknown to Latte in 1925. The fact that the list circulated and was copied in humanistic circles in the later 15th century already indicates a concrete interest in the didactics of Greek dialects at that time, and, consequently, an awareness that Greek appeared in various forms in the literary texts transmitted to the Renaissance.

Unlike other extant ancient sources reflecting on Greek dialects – preserved mainly in Byzantine lexicographical tradition, such as Gregory of Corinth's Περὶ διαλέκτων or Johannes Grammaticus' Compendia Περὶ Αἰολίδος, which approach the dialects primarily in terms of phonetic or morphological differences – this document attempts to assign particular lexical items to specific regions of Greece. This feature alone renders the list noteworthy.

Although the collection may appear unimpressive at first glance, it raises several questions that remain unresolved: What exactly is the nature of the list, and for what purpose was it compiled? What might be inferred regarding its origins and sources? Can any of the forms be plausibly regarded as genuinely ancient or reliably attributed to specific dialects? Finally, to what extent is the list a trustworthy source for dialectological research? Bowra's suggestion – that although the items are poetic, the compiler regards them as current in various local vernaculars (Bowra 1959: 45) – deserves consideration, though it remains difficult to substantiate.

From a strictly dialectological perspective, the list cannot be treated as a dependable source of data. The overwhelming majority of the items are archaic and predominantly poetic. Poetic register appears to have been a decisive factor – of the hundred forms, eighty-five are attested in Homer, though they also occur elsewhere in poetry. Eleven additional forms belong to poetic vocabulary outside traditional Epic diction – appearing in choral lyric, iambus, tragedy, or Hellenistic epic – and most of these have no attestations in prose (Bowra 1959: 46). According to Latte and Bowra, three items appear corrupted through the history of transmission: φύς, given as Doric; a supposed Aeolic κεκρυφάλεος (but see Sowa 2011: 173f.); and the Aetolian οὔλας (Bowra following Latte, ibidem). Of the whole collection, only one form – ἔστιοι, glossed with the puzzling meaning 'dead' and attributed to the Arcadian city of Cleitor (Κλειτορίων ἔστιοι = νεκροί) – lacks any parallels in literary sources. Whether this constitutes genuine evidence for the vernacular of the region remains doubtful.

The entries are arranged according to twenty-two geographical areas, some corresponding to recognised dialect regions, others referring to single cities – an interesting feature in itself.[21] For example, within the Arcadian group, a general dialectal label (Ἀρκάδων) appears alongside separate subgroups assigned specifically to the speech of Cleitor and Phlius. These cities belong, of course, to the Arcadian sphere, but they are not otherwise treated separately in lexicographical sources. A similar situation arises in the case of Argos, which appears both as a dialectal region (Ἀργείων: αἶσα ~ μοῖρα, κτύπος ~ ψόφος, μῆλα ~ πρόβατα) and, independently, through the forms attributed to the Argive city of Hermione, which likewise receives a pair of lexical items.

Κυπρίων τόργος ~ γύψ

In modern scholarship, Cypriot dialectal vocabulary has generally been regarded as belonging to a particularly archaic stratum of Greek, directly continuing a number of Myce-

[21] The list includes the following regions and towns (in the order presented): Athens, Argos, Arcadia, Achaea, Aetolia, Acarnania, Ambracia, Aeolis (referring to Lesbos and the Aeolic cities of Asia Minor), Hermione, Thessalia, Cyprus, Boeotia, Doris, Ionia, Cleitor, Crete, Corinth, Corcyra, Laconia, Magnesia (in Thessaly), Sicily, and Phlius.

naean forms. This has led to the frequent characterisation of the Cypriot dialect – alongside Arcadian – as among the most archaic of the Greek linguistic varieties. Numerous archaic formations attested both in Homer and in Cypriot (or Arcadian) have been traced back to the so-called 'Achaean' layer of Epic diction. According to Ruijgh, this label denotes elements in the language of Homer that predate both the Aeolic and the Ionic strata. On this view, the 'Achaean' dialect represents forms of Greek spoken in the Peloponnese and the adjacent Aegean islands during the Mycenaean period; its remnants persisted into classical times chiefly in Arcadia and Cyprus, while its earliest attestations are to be sought in the Linear B tablets (cf. Ruijgh 1957).

Under the heading 'Cypriots', the anonymous compiler of the Γλῶσσαι includes the form τόργος, glossed as 'γύψ', the name of a bird ('vulture'). The noun is also attested in other lexicographical sources, though occasionally assigned to different dialect regions. Thus, Hesychius T 1161 attributes the form to Sicily: τόργος· εἶδος γυπὸς αἱματορ<ρ>όφου. ἔστι δὲ καὶ ὁ γὺψ παρὰ Σικελιώταις. Without dialectal specification the term appears in Suda T 788, τόργος· παρὰ Λυκόφρονι ὁ κύκνος· λέγεται δὲ καὶ ὁ γύψ, and in Photios T 388, τόργος· ὁ γύψ.

The word is found exclusively in learned Alexandrian poetry. It appears in Callimachus (fr. 647: †καθνώδει τόργος ἔκοπτε νέκυν '*torgos* struck the corpse with its beak') and in Lycophron (*Alexandra* 88): ἣν τόργος ὑγρόφοιτος ἐκλοχεύεται 'the wet-nesting *torgos* broods over it'). The scholia to Lycophron further comment on the semantic development from 'vulture' to 'swan' and the scholiast's elaborate mythological explanation is well known.[22] The glosses, however, reflect literary interpretation rather than living vernacular usage.

The form is exceedingly rare in Greek. It is unattested in the epigraphic material of any region. Consequently, there is no evidence to support the compiler's attribution of the word to the Cypriot dialect (cf. Egetmeyer 2010: 245). From a formal perspective, its etymology is highly uncertain. Attempts to derive it from **tréh₃g/ĝ-* 'to gnaw' are unconvincing on phonological grounds; one would expect a reflex **tr̥h₃g/ĝ- > trōg-* with lengthened vowel (cf. Greek τραγεῖν 'to gnaw'), or a remodelled zero-grade aorist **treh₃g/ĝ-* :: **tr̥h₃g/ĝ-* (LIV² 651; see also DELG s.v.; Hadjioannou 1977: no. 237, van Windekens 1986, Beekes 2010:

[22] Schol. in Lycoph. 88: τόργος ὑγρόφοιτος – ὁ Ζεὺς ἢ ἡ Νέμεσις, παρόσον κύκνῳ ἀπεικασθεὶς ὁ Ζεὺς Νεμέσει τῇ Ὠκεανοῦ συνῆλθεν, ἐξ ἧς γεννᾶται ᾠόν, ὅπερ λαβοῦσα ἡ Λήδα ἐθέρμαινε καὶ ἔτεκε τὴν Ἑλένην καὶ τοὺς Διοσκούρους. τὸ δὲ ὑγρόφοιτος γράφεται καὶ ὑψίφοιτος. τόργος κυρίως ὁ γύψ· νῦν δὲ τὸν κύκνον λέγει, ὃν μιμησάμενος ὁ Ζεὺς συνεμίγη τῇ Λήδᾳ. ὑγρόφοιτος δὲ ὁ ἐν τοῖς ὑγροῖς φοιτῶν καὶ ἀναστρεφόμενος. ὁ γὺψ νῦν δὲ τὸν ἀετὸν ἢ τὸν κύκνον λέγει. ἔλαβε δὲ ζῶον ἀντὶ ζώου. Ζεὺς γὰρ ὁμοιωθεὶς κύκνῳ Νεμέσει τῇ τοῦ Ὠκεανοῦ θυγατρὶ συνῆλθεν εἰς χῆνα, ὡς ληροῦσιν (Ap. III 127), αὐτὴν μεταβαλών· ἡ δὲ τεκοῦσα ᾠὸν ἐν τῷ νικῆς ἢ διότι πολλοὶ δι' αὐτὴν ἀπέθανον. 'The moisture-loving (ὑγρόφοιτος) ***torgos*** – either Zeus or Nemesis, since Zeus, appearing in the form of a swan, united with Nemesis, daughter of Oceanus. From this union an egg was produced, which Leda, having taken it, warmed and gave birth to Helen and the Dioscuri. The term ὑγρόφοιτος is also written ὑψίφοιτος. ***torgos*** properly means the vulture, but now refers to the swan, which Zeus, imitating it, adopted when he united with Leda. 'Moisture-loving' refers to one who moves about and frequents watery places. The word ***gyps*** now also denotes the eagle or the swan. He (the scholiast) has taken 'animal' for another 'animal'. For Zeus, having assumed the form of a swan, united with Nemesis, daughter of Oceanus, as a goose, according to the mythographers (Apollodorus III 127), transforming her; and she, after giving birth, produced an egg on account of which many perished.'

1494). Thompson (1895: s.v.) proposed a Coptic origin (Coptic *t(o)re*, *ere* 'kite'), which remains speculative. A putative connection with Proto-Germanic **sturkaz* (German *Storch*, English *stork*) is likewise highly improbable.

The reasons why such a rare form was selected for inclusion in a supposedly dialectal list remain obscure. What can be stated with certainty is that its designation as specifically Cypriot is untenable.

Σικελῶν κόρσας – κεφαλάς

The forms from Sicily are not very frequent in the ancient lexicographical sources (cf. four glosses of Sicilian origin in the *Lexicon* of Hesychius). It is, therefore, interesting to find words ascribed to this island in the List. One should observe from the outset that the form κόρσας occurs alongside formations such as μέλαθρον· οἰκία, ναίει· οἰκεῖ, or φόρμιγξ· κιθάρα, which doubtless entered the List via Epic poetry (cf. Bowra 1959: 53, on ναίει).

The noun is attested from Homeric times, although in Epic, the Ionic form κόρση is generally preferred. In Attic, the form κόρρη is applied, and in Doric, the related κόρρα is also attested (cf. Theoc. 14.34). Aeolic poetry employed the form κόρσα (cf. Alc. 34: αὐτὰρ ἀμφὶ κόρσαι). The noun has several meanings, of which the glossed 'head' is only one; it may also denote the 'temple' or 'side of the forehead' (e.g., E 584: ξίφει ἤλασε κόρσην), the 'jaw' (especially Attic expressions πατάξαι ἐπὶ κόρρης 'smack on the jaw'; Pherecr. 155b [CAF iii, p. 716]; cf. Plato, *Gorgias* 486c, 508d, 527a: ἐπὶ κόρρης τύπτειν), 'hair' (e.g., Aesch. *Agamemnon* 282: λευκὰς δὲ κ. τῇδ' ἐπαντέλλειν), and 'head' (Empedocles 57.1: κ. ἀναύχενες; cf. Nic. *Th.* 905; Opp. C. 3.25). In Attic, the meaning may extend to the whole head and neck, whereas in Ionic, it refers only to the head.

The form is primarily poetic, with a prose counterpart in κρόταφος, though Attic usage demonstrates possible application in prose. The meaning 'hair' should likely be interpreted metaphorically (Frisk 1960: 923), with a primary semantic sense of 'haircut on the head'. Hesychius, however, lists the various meanings, e.g. 'hair of the eyebrows, jaws, summits, headlocks, ramparts, bulwarks, crowns of towers, temple, staircases' without reference to any particular dialect (cf. K 3660ff.).[23]

There is broad consensus that the form derives from an o-grade *kors-ó-* 'that which is shaven', linked to the verb κείρειν (cf. Hsch. K 3665ff. κορσόν· κορμόν, κορσοῦν· κείρειν) < *(s)ker-* 'to shave' (cf. Alb. *shqerr*; LIV^2: 560f.). However, some forms may derive from the root *kers-* (LIV^2: 358f.), as indicated by the presence of */s/* in compounds such as ἀκερσεκόμης 'with unshaven hair' (DELG: 568, Frisk 1960: 923, Beekes 2010: 755). Since the form exhibits an /o/ vowel via apophony, this cannot serve as proof of its dialectal affiliation, e.g. as an Aeolic or Arcado-Cypriot treatment of the sonorant *r̥*.

[23] **κόρσαι**· αἱ τῶν ὀφρύων τρίχες καταφέρουσαι εἰς τοὺς ὀφθαλμούς. ἢ γνάθοι. ἢ κορυφαί… κεφαλίδας, ἐπάλξεις, προμαχῶνας. στεφάναι πύργων. ἢ κρόταφοι. ἢ κλίμακες; **κόρση**· κεφαλή. ἔπαλξις. κλῖμαξ. κρόταφος. Cf. however *Schol. In Il*. 4, 502a, 502b attributing the meaning 'jaw' to Attic, cf. κόρσην: τὸν κρόταφον· ἐπὶ κόρσης γὰρ λέγουσιν Ἀττικοὶ τὴν γνάθον.

Considering internal developments of Greek, the form exhibits no distinct Doric features, as might be expected for Sicily,[24] and the Doric variant κόρρα appears in Theocritus (14.34: τᾶμος ἐγώ, τὸν ἴσαις τύ, Θυώνιχε, πὺξ ἐπὶ κόρρας). The alternative form κόρσης occurs in the work of the same author (25.255: κόρσης ὕπερ αὖον ἀείρας) likely for stylistic variation to avoid repeating κεφαλή.[25] The forms κόρρη or κόρση are semantically indistinguishable from κρόταφος (Pollux 2.40*: κράνιον*, with some calling κροτάφους or κόρρας; Rufinus, *Eph. Onom.* 13: τὰ δὲ ἑκατέρωθεν τοῦ βρέγματος κόρσαι καὶ κρόταφοι). The expression ἐπί κόρρης denotes striking someone's ears with the flat of the hand (Demosth. 21.72).

It is difficult to decide which form should be considered more 'authentic' in Theocritus' literary dialect. Conversely, the nominative plural κόρσαι in Alcaeus should be treated as Aeolic. The cluster /rs/ cannot be assigned to the literary Lesbian dialect. In Aeolic, *-Vrs-* generally develops into a short vowel plus a geminated sonorant, whereas other dialects yield a long vowel plus a single sonorant (cf. Lesb. στελλειν vs. non-Aeolic στειλειν; Lesb. ορρανος vs. non-Aeolic ουρανος). In literary Lesbian, further simplification of geminated resonants for metrical reasons also occurs (Blümel 1982: 102). Forms retaining /rs/ and /ls/ are widely spread in certain words (e.g., Hom. κέλσαι, ἔκερσεν, ἄρσην, and ἔρσην), whereas in some dialects the cluster /rs/ assimilated to /rr/ without affecting the preceding vowel (Attic ἄρρην, θάρρος), sometimes under the influence of Ionic. Similar phenomena occur in Western Ionic and Arcadian (e.g., φθεραι < φθερσαι; Lycophr.), Elean, Doric (Thera, Delphi), and authors such as Alcman, Epicharmus, Sophron, and Plutarch. Even in dialects that regularly have /rr/, original */rs/* may be retained by analogy (e.g., Attic καθάρσις), reflecting Ionic (Homeric) or later Koine influence.

Thus, the form κόρσας in the List, allegedly Sicilian, likely represents a form retaining the consonant cluster. It is plausible that it is in fact Ionic or Homeric, erroneously ascribed to Sicily. While a memory of Ionic presence in Sicily may have survived in grammatical tradition, this is unlikely, given that Ionic usage in Chalcidian colonies (Zancle, Himera, Leontinoi, etc.) ceased after the fifth century BCE, with only Doric traces thereafter. The true rationale for classifying κόρσας as Sicilian remains obscure. It is possible that the form κόρση was used by Empedocles (B 57: κόρσαι ἀναύχενες ἐβλάστησαν) and became associated with him (e.g. ἀναύχενος κόρση and other phrases attributed to Empedocles, Simplicius in *Cael.* 586.30; cf. Cat. 337.2;[26] Joh. Philoponus, *In libros de generatione animalium* 14.3,

[24] Cf. Mimbrera (2012: 191-222) for the description of dialectal situation in Sicily in Classical period.

[25] Scholia in Theocr. 14, 34a 2 κόρρην γὰρ τὴν γνάθον καὶ σιαγόνα οἱ Ἀττικοί. <κόρρας:> τὰς σιαγόνας ἢ τὰς μήνιγγας

[26] Arist. De caelo Γ 2. 300b 25

ἦι πολλαὶ μὲν κόρσαι ἀναύχενες ἐβλάστησαν

γυμνοὶ δ' ἐπλάζοντο βραχίονες εὔνιδες ὤμων,

ὄμματά τ' οἶ(α) ἐπλανᾶτο πενητεύοντα μετώπων.

'On it (the earth) many heads sprung up without necks and arms wandered bare and bereft of shoulders. Eyes strayed up and down in want of foreheads' R. P. 173 a.'; cf. also Simpl. *De caelo* 586, 29 ἂν εἴη μίξεως σημαντικὸν ἡ ἀναύχενος κόρση καὶ τἆλλα τὰ ὑπὸ τοῦ Ἐμπεδοκλέους λεγόμενα 'γυμνοὶ ... μετώπων' καὶ πολλὰ ἄλλα, ἅπερ οὐκ ἔστι μίξεως παραδείγματα.

27, 31ff.: τὰς ἀναύχενας κόρσας).[27] Given Empedocles' Sicilian origin, the Homeric/Ionic form he used may have been mistakenly associated with Sicily as the ultimate source.

***

Setting aside the dialectological assessment of individual items, the principal issue concerns the rationale behind the Compiler's ordering of forms and, more broadly, the relationship between the List and the lexicon of Hesychius – the most substantial surviving lexicographical source. All items in the List (with the exception of οὔλας and ἔστιοι) are found in Hesychius, though without dialectal labels. As argued elsewhere (Sowa 2011: 172-179), these labels reflect neither linguistic reality nor vernacular usage. They derive instead from Homeric diction and have been arbitrarily, even randomly, assigned to dialects by the Compiler. This does not preclude the possibility that certain words are genuinely dialectal, but if so, their dialectal affiliation seldom corresponds to that suggested in the List (Sowa 2023: 122 ff.). Taken together, the analysis of the List and its relationship to Gregory of Corinth illustrates that ancient conceptions of dialectal 'normativity' were literary and prescriptive rather than descriptive in the modern sense. Both works illuminate ancient conceptions of Greek dialects, showing how literary and scholarly authority, rather than spoken practice, shaped perceptions of dialect correctness. The Compiler's assignment of forms to dialects reflects an interpretive tradition rooted in poetic and scholarly authority, rather than observation of vernacular usage. This reinforces the need to distinguish between ancient notions of dialect correctness and modern dialectological methodology.

The works of Gregory of Corinth and the anonymous Γλῶσσαι are fundamentally different, yet both illuminate ancient conceptions of Greek dialects. Returning to the initial question – whether a form of normativity can be posited in the lexical systems of Ancient Greek dialects – one must distinguish clearly between the ancient grammatical tradition and modern dialectological analysis. From the modern perspective, dialectal texts exhibit highly heterogeneous forms: vernacular dialectal items intermingle with literary or poetic usage, particularly in metrical funerary inscriptions; obsolete forms are replaced by newer or more frequent ones, or by forms belonging to a prestigious dialect (e.g. Attic or the κοινή). Greek dialects undoubtedly possessed social varieties, like any language, though the epigraphic record seldom allows such nuances to be recovered (García Ramón 2018: 64).

Linguistic variation continued to be conspicuous well into the Classical and Hellenistic periods, as inscriptions with vernacular features demonstrate. Literary practice also shaped contemporary expectations – Homer, studied intensively in education, familiarised

[27] ἀλλὰ καὶ τὰ ὑπὸ τοῦ Ἐμπεδοκλέους λεγόμενα ἀδύνατά ἐστι, κἂν συμφωνότερα τοῖς φαινομένοις τυγχάνῃ. καὶ πῶς ἀδύνατα, ἐπάγει· ὥσπερ γάρ, φησί, καὶ μεγάλα ὄντα τὰ ἀνομοιομερῆ, ἡνίκα τὸ νεῖκος τῆς φιλίας ἐπεκράτησεν, οὐκ ἠδύνατο ἔμψυχα εἶναι (ἐμψύχους γὰρ καὶ αἴσθησιν ἐχούσας ἔλεγε τὰς ἀναύχενας κόρσας) ὥσπερ οὖν αἱ μεγάλαι ἐκεῖναι κεφαλαὶ οὐκ ἠδύναντο ἔμψυχοι εἶναι 'Even what Empedocles says is impossible, even if it seems more in line with appearances. How so? He explains that large, heterogeneous things, when dominated by the strife of friendship, cannot be living beings. For example, he claimed that the "neckless ***kórsai***" had life and sensation – but in reality, just as those enormous heads could not be alive, these too could not truly possess life or feeling.'

Athenians with non-Attic forms; the Doric of tragic choruses and mélos exposed audiences to further dialect features. Phonological contrasts such as the Ionic-Attic shift /a:/ → /e:/ or dialectal pronominal forms were readily perceptible to such educated readers. Only local slang and specialised vocabulary may have posed difficulty.

Gregory of Corinth's approach to dialectal 'normativity' is neither rigid nor absolute; he does not enforce a single standard but provides criteria for choosing among alternatives. He acknowledges and explains linguistic diversity rather than suppressing it and treats normativity as a tool for clarity and stylistic refinement. By contrast, the Γλῶσσαι do not refer to any standard form; instead, they consistently classify forms rooted in poetic diction as dialectal. As with modern languages, a dialect may constitute a complete linguistic system, but in Ancient Greek the high degree of shared material makes it difficult to determine what constitutes a specifically dialectal item, what merits lexicographical inclusion, and what is 'non-standard' relative to a presumed norm. In effect, the entire corpus is 'non-standard', and lexicographical treatment must therefore accommodate interdialectal influence, stylistic or functional variation, and chronological stratification. Ancient grammatical and lexicographical sources largely treat Homer as the repository of all non-standard Greek forms. By the time of Diogenianus, Hesychius, and Kyrillos – prior to the Atticist movement and its purist ideology – the dialects appear to have become an abstract construct – a storage place for all rare, obsolete, or otherwise anomalous words.

References

The Lesbian inscriptions are quoted by a number after Hodot 1990, pp. 272-317 using the specification of the finding place of Lesbian inscriptions as: MYT Mytilene, MAT Methymna, ERE Eressos, LES Lesbos without determination of a city and NAS the island of Nessos. The ancient authors have been quoted according to the electronic editions in Thesaurus Linguae Graecae Project. The abbreviations follow the model as in LSJ. The Lesbian poets, however, have been quoted according to the edition of E. Lobel and D. L. Page, *Poetarum Lesbiorum fragmenta*, Oxford, Clarendon Press, 1955, and according to D. L. Page, *Supplementum lyricis Graecis*, Oxford, Clarendon Press, 1974. The non-Aeolic Inscriptions and Papyri have been quoted after the electronic edition in PHI CD 6/7 (prepared by the Greek Epigraphy Project at Cornell University). The abbreviations follow the standard model.

Adams, D. Q. 2013. *A dictionary of Tocharian B, revised and greatly enlarged*. Amsterdam – New York: Rodopi.

Becares, V. 1988. Ein unbekanntes Werk des Gregorios von Korinth und seine Lebenszeit. *Byzantinische Zeitschrift* 81. 247-248.

Bechtel, F. 1921. *Die griechischen Dialekte.* Erster Band: *Der lesbische, thessalische, böotische, arkadische und kyprische Dialekt*. Berlin: Weidmannsche Buchhandlung.

Beekes, R. 2010. *Etymological dictionary of Greek* (with the assistance of L. van Beek). Boston – Leiden: Brill.

Bekker, I. 1821. *Anecdota Graeca III*. Berlin: Typis et impensis G. Reimeri.

Blümel, W. 1982. *Die aiolischen Dialekte: Phonologie und Morphologie der inschriftlichen Texte aus generativer Sicht*. Göttingen: Vandenhoeck & Ruprecht.

Bolognesi, G. 1953. Sul Περὶ διαλέκτων di Gregorio di Corinto. *Aevum* 27. 97-120.

Bonner, J. 1909. The mutual intelligibility of Greek dialects. *The Classical Journal* 4. 356-363.

Bowra, C. M. 1959. Γλῶσσαι κατά πόλεις. *Glotta* 38(1-2). 43-60.

Brandenburg, P. 2005. *Apollonios Dyskolos: Über das Pronomen: Einführung, Text, Übersetzung und Erläuterungen*. München – Leipzig: K. G. Saur.

Buck, C. D. 1955. *The Greek dialects: grammar, selected inscriptions, glossary*, Chicago.

Chantraine, P. 1948. *Grammaire homérique. Tome I : Phonétique et morphologie*. Paris: Libraire C. Klincksieck.
Colvin, S. 1999. *Dialect in Aristophanes: The politics of language in Ancient Greek literature*. Oxford: Clarendon Press.
Coo, L. 2021. Sappho in fifth- and fourth-century Greek literature. In Finglass, P. J. & Kelly, A. (eds.), *The Cambridge companion to Sappho*, 263-276. Cambridge: Cambridge University Press.
de Andres, G. 1986. *Catálogo de los códices griegos de la Biblioteca Nacional*. Madrid: Ministerio de Cultura Dirección General del Libro y Bibliotecas.
DELG = Chantraine, P. *Dictionaire etymologique de la langue grecque: Histoire des mots*. Tome I: *(A-D)* 1968, Tome II: *(E-K)* 1970, Tome III: *(L-P)* 1974. Paris: Klincksieck.
Dickey, E. 2007. *Ancient Greek scholarship: A guide to finding, reading and understanding scholia, commentaries, lexica and grammatical treatises, from their beginnings to the Byzantine Period*. Oxford: Oxford University Press.
Donnet, D. 1966. Précisions sur les oeuvres profanes de Grégoire de Corinthe. *Bulletin de l'Institut Historique Belge de Rome* 37. 81-97.
Donnet, D. 1967. *Le traité Peri suntaxeōs logou de Grégoire de Corinthe: Étude de la tradition manuscrite*. Bruxelles – Rome: L'Institut historique belge de Rome.
Dubois, L. 1986. *Recherches sur le dialecte arcadien.* I. *Grammaire.* II. *Corpus dialectal.* III. *Notes – index – bibliographie.* Louvain-La-Neuve: Cabay.
Dunkel, G. E. 2004. Particles and personal pronouns: Inclusive **me* and exclusive **μe*. In Penney, J. H. W. (ed.), *Indo-European perspectives: Studies in honour of Anna Morpurgo Davies*, 18-29. Oxford: Oxford University Press.
Egetmeyer, M. 2010. *Le dialecte grec ancien de Chypre*. Berlin – New York: de Gruyter.
Frisk, H. 1960. *Griechisches Etymologisches Wörterbuch*. Heidelberg: C. Winter..
García Ramón, J. L. 1975. *Les origines postmycéniennes du groupe dialectal éolien*. Salamanca: Universidad de Salamanca.
García Ramón, J. L. 1999. Cuestiones de léxico y onomástica tessalios. In Cassio, A. C. (ed.), *Katà diálekton: Atti del III Colloquio Internazionale di Dialettologia Greca – Napoli – Fiaiano d`Ischia, 25-28 settembre 1996*, 521-552. Napoli: Istituto Universitario Orientale.
García Ramón, J. L. 2018. Ancient Greek dialectology: Old and new questions, recent developments. In Giannakis, G. & Crespo, E. & Filos, P. (eds.), *Studies in Ancient Greek dialects from Central Greece to the Black Sea*, 29-106. Berlin: de Gruyter.
Glucker, J. 1970. Thucydides I 29, 3, Gregory of Corinth and the *ars interpretandi*. *Mnemosyne* 4th series. 23. 127-149.
Hadjioannou, K. P. 1977. *Ἡ ἀρχάια Κύπρος εἰς τὰς ἑλληνικὰς πηγάς.* Τόμος 3.2: *Κυπρίων Γλῶσσαι*. Nicosie: Ἔκδοσις ἱερᾶς ἀρχιεπισκοπῆς Κύπρου.
Hamm, E.-M. 1957. *Grammatik zu Sappho und Alkaios*. Berlin: de Gruyter.
Hernández Vázquez, A. 1994. *Estudio léxico del jonio minorasiático*. Salamanca: Universidad de Salamanca. (Doctoral dissertation.)
Hesych. = *Hesychii Alexandrini Lexicon*. Vol. 1 A-Δ ed. Latte, K. & Cunningham, I. C. 2018; Vol. II E-O ed. Latte, K. & Cunningham, I. C. 2020; Vol. III Π-Σ ed. Hansen, P. A. 2005; Vol. IV T-Ω ed. Hanses, P. A. & Cunningham, I. C. 2009; Berlin – New York: Walter de Gruyter (Sammlung Griechischer und Lateinischer Grammatiker 11, 1-4).
Hodot, R. 1990. *Le dialecte éolien d'Asie: La langue des inscriptions viie s. a.C. ive s. p.C.* Paris: Éditions Recherche sur les civilisations.
Hoffmann, O. 1893. *Die griechischen Dialekte in ihrem historischen Zusammenhange.* 2. Band: *Der nord--achäische Dialekt*. Göttingen: Vandenhoeck und Ruprecht.
Hunger, H. 1982. Gregorios von Korinth, Epigramme auf die Feste des Dodekaorton. *Analecta Bollandiana* 100. 637-651.
Iriarte, J. 1769. *Regiae bibliothecae matritensis codices graeci mss*. Madrid.
Kloekhorst, A. 2008. *Etymological dictionary of the Hittite inherited lexicon*. Leiden – Boston: Brill.
Kominis, A. 1960. *Gregorio Pardos, metropolita di Corinto, e la sua opera* (Testi e Studi Bizantino-neoellenici 2). Rome – Athens.
Latte, K. 1925. Glossographika. *Philologus* 80. 136-175.
Laurent, V. 1963. Review of Kominis (1960). *Revue des études byzantines* 21. 290-291.

Lejeune, M. 1972. *Phonétique historique du mycénien et du grec ancien*. Paris: Klincksieck.
LIV^2 = Rix, H. & Kümmel, M. & Zehnder T. & Lipp, R. & Schirmer, B. (eds.) 2001. *Lexikon der indogermanischen Verben: Die Wurzeln und ihren Primärstammbildungen*. 2nd edn. Wiesbaden: Reichert.
Matzinger, J. 2012. 'Zwischensprachen' – Areallinguistische Bemerkungen aus dem Bereich des Balkanindogermanischen. In Sadovski, V. & Stifter, D. (eds.), *Iranistische und indogermanistische Beiträge in memoriam Jochem Schindler (1944-1994)*, 137-159. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
Meier-Brügger, M. 1986. Homerische Kunstsprache: Synchron und diachron. *Glotta* 64. 127-143.
Miller, D. G. 2014. *Ancient Greek dialects and early authors: Introduction to the dialect mixture in Homer, with notes on lyric and Herodotus*. Berlin: de Gruyter.
Mimbrera, S. 2012. Sicilian Greek before the fourth century BCE. In Tribulato, O. (ed.), *Language and linguistic contact in Ancient Sicily*, 191-222. Cambridge: Cambridge University Press.
Montana, F. 1995. *Gregorio di Corinto: Esegesi al Canone Giambico per la Pentecoste attribuito a Giovanni Damasceno*. Pisa: Giardini.
Morpurgo Davies, A. 1968. Article and demonstrative: A note. *Glotta* 46. 76-85.
Morpurgo Davies, A. 1987. The Greek notion of dialect. *Verbum* 10. 7-28.
OED = Simpson, J. A. & Weiner, E. S. C. 1989. *The Oxford English dictionary*. 2nd edn. Vol. 4. Oxford: Clarendon Press.
Parker, H. N. 2008. The linguistic case for the Aiolian migration reconsidered. *Hesperia* 77. 431-464.
Peters, M. 1980. *Untersuchungen zur Vertretung der Indogermanischen Laryngale im Griechischen*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
Peters, M. 1986. Zur Frage einer "achäischen" Phase des griechischen Epos. In Etter, A. (ed.), *O-o-pe-ro-si: Festschrift für Ernst Risch zum 75. Geburtstag*, 304-319. Berlin – New York: de Gruyter.
Peters, M. 1994. Griech. γῆ, γαῖα, armen. *erkir* 'Erde'. In Vavroušek, P. (ed.)*, Iranian and Indo-European studies: Memorial volume of Otakar Klíma*, 203-213. Praha: Enigma Corporation.
Puhvel, J. 1984- . *Hittite etymological dictionary.* Vol. 1: *Words beginning with A*. Vol. 2: *Words beginning with E and I*. Berlin – New York – Amsterdam: Mouton.
Risch, E. 1974. *Wortbildung der homerischen Sprache*. 2nd edn. Berlin – New York: de Gruyter.
Rix, H. 1992. *Historische Grammatik der griechischen: Laut- und Formenlehre*. 2nd edn. Darmstadt: Wissenschaftliche Buchgesellschaft.
Robins, R. H. 1993. *The Byzantine grammarians: Their place in history*. Berlin – New York: Mouton de Gruyter.
Rodríguez Somolinos, H. 1998. *El léxico de los poetas lesbios*. Madrid: Instituto de Filología – Consejo Superior de Investigaciones Científicas.
Ruijgh, C. J. 1957. *L'élément achéen dans la langue épique*. Assen: Van Gorcum.
Ruijgh, C. J. 1971. *Autour de 'τε épique'*. Amsterdam: Adolf M. Hakkert.
Scarborough, M. 2023. *The Aeolic dialects of Ancient Greek: A study in historical dialectology and linguistic classification*. Leiden – Boston: Brill.
Schäfer, G. H. (ed.). 1811. *Gregorii Corinthii et aliorum grammaticorum libri de dialectis linguae graecae, quibus additur nunc primus editus Manuelis Moschopuli libellus de vocum passionibus*. Leipzig: A. G. Weigel
Schmidt, J. 1900. Das Zahlwort μία, ἓα. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 36. 391-399.
Schoubben, N. 2019. *Byzantine descriptive linguistics: Introduction, translation and commentary of Gregory of Corinth on Aeolic.* Leiden: University of Leiden. (Seminar paper.)
Sowa, W. 2011. Griechische Dialekte, dialektale Glossen und die antike lexikographische Tradition. *Glotta* 87. 159-183.
Sowa, W. 2023. Γλῶσσαι κατὰ πόλεις: Greek dialects through the lenses of the ancient lexicography. In Machajdíková, B. & Eliášová Buzássyová, Ľ. (eds.), *Greek – Latin – Slavic: Aspects of linguistics and grammatography*, 115-144. Tübingen: Narr Francke Attempto.
Sowa, W. 2024. *Αἰολικά: Studia nad dialektem eolskim na podstawie Περὶ Αἰολίδος Grzegorza z Koryntu*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
Thompson, D. W. 1895. *A glossary of Greek birds*, Oxford: Clarendon Press.
Thumb, A. & Scherer, A. 1959. *Handbuch der griechischen Dialekte.* Bd. II. Heidelberg: Carl Winter Universitätsverlag.

Tribulato, O. 2019. Making the case for a linguistic investigation of Greek lexicography: Some examples from the Byzantine reception of Atticist lemmas. In Passa, E. & Tribulato, O. (eds.), *The paths of Greek: Literature, linguistics and epigraphy*, 241-270. Berlin – Boston: de Gruyter.

Tucker, E. F. 1990. *The creation of morphological regularity: Early Greek verbs in* -éō*,* -áō*,* -óō*,* -úō*, and* -íō. Göttingen: Vandenhoeck & Ruprecht.

van Windekens, A. J. 1952. *Le Pélasgique: Essai sur une langue indo-européene préhellénique*. Louvain: Institut Orientaliste.

Wilson, N. G. 1996. *Scholars of Byzantium*. London: Duckworth & Co. Ltd.

Wyatt, W. F. 1992. Homer's linguistic forebears. *Journal of Hellenic Studies* 112. 167-173.

Yatromanalakis, D. 2007. *Sappho in the making: The early reception*. Washington: Centre for Hellenic Studies.

DOI: 10.14746/linpo.2025.67.2.5

Carmen Dagostino, Marianne Mithun, and Karen Rice (eds.) 2023 (vol. 1), 2024 (vol. 2). *The languages and linguistics of indigenous North America: A comprehensive guide, Volume 1; Volume 2*. (The World of Linguistics 13.1-2). Berlin – Boston: De Gruyter Mouton. Pp. LII (front matter) + 1-715 + separate enlarged map insert (vol. 1), and + pp. XII (front matter) + [987] 717-1702 (vol. 2)

Alfred F. Majewicz

International Institute of Ethnolinguistic and Oriental Studies, Stęszew
majewicz@amu.edu.pl | ORCID: 0000-0002-8984-3148

... listing Nez Perce *words*
[...] is like writing a dictionary of *sentences* in English.
(Aoki 1994: x)

This reviewer has since his early childhood been interested "in languages" but, as he recalls now, it was a chance encounter with two bulky volumes entitled *Handbook of American Indian Languages* (Boas 1911, 1922), followed by an outburst of utmost fascination with descriptions of tongues with unimaginable before both phonetic as well as morphological structures[1], that resulted in the decision to make linguistics his profession. To be sure on

[1] The languages fairly extensively described being Athapascan (Hupa), Tlingit, Haida, Tsimshian, Kwakiutl, Chinook, Maidu, Algonquian (Fox), Siouan – Dakota (Teton and Santee dialects, with remarks on the Ponca and Winnebago), Eskimo in vol. 1 (1911), and Takelma, Coos, Siuskavan (Upper Umpqua), Chukchee in vol. 2 (1922). Eskimo and Chukchee seemed out-of-the-place in the set as not necessarily American, the former being transborder and transcontinental (in use from Canada westwards to Alaska and Russian Chukotka (Asia), and eastwards, via Labrador on the way, to Greenland (autonomous territory of the Kingdom of Denmark, Europe), the latter in use in Chukotka (Russia) but Boas explained the insertion of the latter (and actually both) in the following way: "It seemed important to add the Chukchee to the sketches contained in the Handbook, because it proves conclusively that those features which are most characteristic of many American languages are found also on the Asiatic continent. It seemed essential, furthermore, to present material for determining the position of the Eskimo language in relation to all its neighbors" (1922:637). We shall come back to "Eskimo" in this text toward its end.

LP LXVII (2), 2025. © The Author(s). This is an open access article licensed under the CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en/).

obvious reasons, this writer never intended to become a researcher or "specialist" in Amerindian languages[2] but both "Boas volumes" served for several decades as one of sources of linguistic data frequently referred to, or used as illustrative examples, in university courses and lectures, and conference presentations.

Now, the situation is different, much better: extensive research brought numerous modern descriptions, grammars, dictionaries of languages insufficiently, partially, or never documented which revealed abundant linguistic facts and phenomena even more attractive for such purposes. Nevertheless, works which competently sum up and generalize, like for this case, Goddard 1996[3], Campbell & Mithun 1979, Campbell 1997, Mithun 1999, or the work under concern here, are awaited, looked for, and welcomed on the linguistic market.

The "Preface" (cf. below) categorizes the publication as a *handbook* and this author will adapt this handy term in the present text.

The fairly extensive front matter (52 pp.) includes a two-page "Preface" (V-VI), six-page "Table of Contents" (VII-XII; these two components have been repeated as front matter in vol. 2), eleven maps with a special "Introduction" with "References" to them (XLI-LII), and the, dominant in this part of vol. 1 (XIII-XL), extremely useful "List of North American families, languages, and dialects" arranged alphabetically and cleverly tabularized, with the head entries in the left-side column listing glottonyms for all the three classificatory level items, provided in the parallel middle column with "alternate names and spellings" (!), and with the information, whenever appropriate, on the "family (branch)" affiliation[4]. Maps appear also on the pages of individual texts throughout the edition (pp. 269 and 271 (morphological types and variation of negatives), 566 (Ahtna speaker's 1912 trip map), 570 (a screenshot locating Native Land), 671 (John Powel's 1891 map of 58 language families), 672 (Sapir's macro-families), 674 (Voegelins map of 1965, phyla and families), 931 (Algic), 1014 (Wakashan), 1116 (Kiksht, Chinook), 1142 (Sahaptian), 1170 (Karuk), 1253 (Californian – "Key to Tribal Territories"), 1305 (Yuman), 1520 (Chitimacha), 1628 (location of unclassified extinct languages))[5]. A 1999 "revised and enlarged [70 (width)×64 cm separate sheet], with additions and corrections" map of "Native languages and language families of North America" by Ives Goddard (1996) is attached in vol. 1; it verbally locates 452 (rough count) glottonyms ("linguistic units"), including areas of 34 language families and 28 individual languages distinguished by numbers (1-62) and color shades. There are also many other illustrations in the two books.

[2] with years passing, focusing primarily on minor, "lesser-used" tongues (*langues moins répandues*) of the Far East, some structurally polysynthetic included.

[3] Including twelve grammatical sketches (of Central Alaskan Yupik, Hupa, Cree, Lakhota, Zuni, Eastern Pomo, Seneca, Wichita, Thompson, Coahuilteco, Sahaptin, and Shoshone).

[4] This "List" alone, as well as the eleven maps, if one cannot afford having access (or for whom such access is not "a must") to the handbook, are publicly available in pdf.

[5] This writer had no intention to make this recital exhaustive but probably it is complete (there is no special list of maps for the entire edition, so it can prove useful).

The core of the *handbook* is organized into eight units (let me call them "parts") marked with Roman numerals I-VIII each in turn divided into chapters marked with Arabic numerals 1-61) authored by 82 contributors, among them also native users of individual languages (their "who is who" provided on pp. 1649-71). Actually, the editors (publishers?) see it in a slightly different way: "The volume [*sic!*] is divided into two main parts, the first on general topics, and the second on revitalization and sketches of languages and families" (p. V in both volumes). Physically vol. 2 (save the front matter) in fact and every aspect is a direct continuation of the physical vol. 1, so this writer also tends to treat the entire work as oneness, perceiving the "part" units his own way. Each chapter is further divided into sections and subsections and ends with its own list of "references", at times with addition of relevant literature not cited in the respective text.

"Part" I (1-179, ***chapters 1-7***) concentrates on characterizing phenomena related to the sound and sounds of the languages, often drastically different from what the "Western world" is familiar with or accustomed to, discussing acoustics and articulations, tone systems, sound and phoneme inventories, and prosody ("word" and "beyond the word"). There are some very informative and useful figures (e.g. IPA and NAPA ("North American Phonetic Alphabet"), convertion charts, tone notations, syllable structures, prosodic morphology (prosodic features as vehicles of categorial~grammatical meanings), phenomena like "lexical tone beyond the word" or "stress beyond the word", intonation).

"Part" II (181-244) opening with a chapter (***8***) attempting to answer the fundamental question "What is a word from the perspective of Indigenous North American languages (183)?" and mentioning "the areas where we need further research to learn more about complex and heterogeneous phenomena related to *word(-like)* units" (*ib.*, *cursive* afm). No wonder, therefore, that the (only) other chapter (***9***) is conceived as "an introduction to word classes" in these languages which "have a unique part to play in research on word classes" (*ib.*)[6].

"Part" III (245-381, seven chapters ***10-16***) aims at elucidating how to put such words into clauses (word order, ergative and nominative-accusative constructions, agreement, negatives, interrogation and requests (immediate and delayed, affirmative and negative, prohibitatives), imperative-only lexemes, information structure[7], focus and topic, case marking, polysynthesis) and clauses into sentences ("clause combining" with "some tricky cases", 323-62) which leads us to:

[6] Try, Dear Reader, to cope with e.g. *igamsiqayugvikumanginaghyaghqaqsaghaghpesikut* (848, quoted from de Reuse 1994: 83) concerning 'thankfulness', or *ayagciqsugnarqnillruuq* informing that 'he said he would probably go' (*ib.*, 196) < *ayag-* 'to leave, to go away, to depart' (Jacobson 2012: 162) and cf. with e.g. *ayagcecissuun* 'starter of an engine' or *ayagcetaag* 'missionary' (*ib.*, 163); you are also encouraged to look for Rubtsova 1971 and find in it a 45pp. (610-44) appendix constituting a list of 764 examples of derivatives of *қимўхси-* 'sled in a dog-team', like for instance *қимўхсиӈнáқутúяхтўғақуқ* 'he intends to hire (somewhere) a dog-team sled for himself' (or, 'he intends to hire (somewhere) a dog-team sled with a driver for himself'; *ib.*, 643). Hence – validity of the above question: indeed, *what is a word*?

[7] – with questions like "what is information structure?" or "how does one talk about information structure in languages with sentences that frequently consist of a single word, as in Unangam Tunuu [Aleut ...] *aniqduĝikuqing* 'I have a child'?" (306-7).

"Part" IV labeled "Discourse", again with two (***17-18***) chapters only: "Verbal art" (385--419, with extensive bibliography of "Further references of interest", 411-9, following "References" 407-11) and "Conversation structure" (421-49).

What comes next is "Part" V "Meaning" (451-616) with eight subsequent chapters (***19-26***) on "Lexicalization and lexical meaning" (453-77), "Lexicography" (a key issue for compiling dictionaries of languages so different from what most linguists have experienced; 479--95), "Evidentiality" (497-510), "Pluractionality and distributivity" (511-26), "Mass and count nouns" (527-46), "Space, landscape, and orientation" (547-76), "A sense of time and world" (547-98: tense(s and tenselessness), 578-82, aspect(s), 583-7, modality and mood, 587-93; references complemented with "other readings of interest"; 577-98), and "Pragmatics" (language and context interactions: "conversational implicatures", 601-6, politeness, 606-9, presuppositions, 609-12, "importance of pragmatics for documentation and revitalization", 612-4).

"Part" VI "Languages over space and time" (617-715), final in vol. 1, with five chapters (***27-31***) on "How grammar can emerge", 619-46, "Language contact and linguistic areas" (*i.a.*, native-native and native-European contacts and their results, mixed languages as "extreme linguistic results" of such contacts), 647-68, "Language classification", 668-87 (including "An abridged history of language classification in North America", 670-5), "Archival-based sociolinguistic variation" (linguistic data retrieval from archival records and their relevance, 689-700), and "Community-based sociolinguistic variation" (701-15).

Part VII "Language revitalization" (719-839) with six chapters (***32-37***) devoted to strategies, methods, problems, resources, etc., conceived and implemented to save from extinction, revive, preserve for generations to come, indigenous North American languages, most of them being seriously-to-critically endangered, or to reclaim those no longer used (dormant or even extinct but recorded in the past and in some petrified way – like written documents or wax cylinder audiorecordings – preserved in museums, libraries, research institutions, ... Consecutive chapters discuss "outcome of a Mentor-Apprentice program/style (MAP) learning" (719-39), first-language acquisition ("child and child-directed speech" in indigenous languages), reviewing published research results (741-66), "pedagogies of decolonization" of these languages (767-88), "digital tools for language revitalization" (789-805), "using archival materials for language reclamation" (807-21)[8], and "*changing* [from "linguist--centred" to "community-centred"] notions of fieldwork" (823-39, *italics* afm).

"Part" VIII "Language families and isolates", the most extensive in the handbook (841--1647), embraces 23 (***38-60***) chapters – sketches of particular language families (19) or language isolates (4) and one chapter (***61***) on extinct "unclassified languages", providing basic data on genetic ties, location, state of preservation, characteristic and unique features in phonetics, phonology, morphology, with numerous illustrative examples:

[8] of special interest to this writer with his years of experience gained while reconstructing Bronisław Piłsudski's results of research on Sakhalin, Hokkaido, and Lower Amur region indigenous peoples and languages (Nivhgu, Ainu, Orok, Ulcha, and Nanai) between 1892 and 1906 and in 1910 (CWBP 1998-2011).

chapter 38 "Inuit-Yupik-Unangan: An overview of the language family" by Richard Compton[9], 843-73; we shall come back to this chapter at the end of this review.

39 "Dene-Athabaskan" by Leslie Saxon (875-930), family including "some 40 languages, and [their] varieties" (no special easy-to-find language list or data on preservation or endangerment of these languages (pity)[10], instead a fairly long and rich list of references (fortunately) provided;

40 "Algonquian" by Will Oxford (931-50), family (or part of Algic family, if Yurok and Wiyot added) with 25-30 languages (more familiar glottonyms~ ethnonyms being Arapaho, Blackfoot, Cheyenne, Cree, Delaware (Lenape), Penobscot (Abenaki), Fox (Meskwaki), Mahikan, Menominee, Micmac (Mi'kmaq), Montagnais (Innu), (Maliseet-)Passamaquoddy, Potawatomi, Shawnee; Cree-Innu-Naskapi dialect cluster with 86,475 (Cree) + 11,605 (Innu--Nascapi) users in 2021 (*Statistics Canada, Census of Population*) is said to constitute the aboriginal language with the highest number of speakers in Canada[11];

41 "Michif" by Nicole Rosen (951-84) with 11 "alternate names and spellings" (XXVII), a hybrid (mixed) language or (glotto~topo)-lect cluster[12] "developed at the turn of 19th century [...] spoken today by likely fewer than a couple hundred [Metis] people in Western Canada and North Dakota" (951);

42 "Tsimshianic" by Clarissa Forbes (985-1012), family consisting of two <Maritime> (Coastal~Sm'algyax, the best known, and Southern~Sgüüxs, no longer spoken, Tsimshian) and two <Interior> (Nisga'a and Gitksan) languages, all treated also as a dialect continuum and seriously endangered, with small and diminishing number of speakers (low hundreds but also between dozens and zero) but "hundreds of self-reported active learners" (986); "other relevant literature" list added[13] (1010-12);

[9] Technical reasons (limited space and character/genre of this text – a *review*) prevented this reviewer from planned providing names of all contributors of all the handbook chapters but he convincingly (reference needs) insisted on identifying contributors for this part.

[10] Among languages better known to linguists-non-specialists in Amerindian tongues are Slave, Chipewyan, Dogrib, Beaver, Carrier, Tutchone, Kutchin, Koyukon, Upper Kuskwokwim, Tanana, Han, Hupa, Wailaki, Tolowa, Navajo (with some 170, 000 – the largest number of speakers of any Amerindian languages north of Mexico), Kiowa Apache, San Carlos (Western) Apache (earlier literature widens the family to include also Eyak and Tlingit but the affinity is considered "an open question"). Slave (a language or language~dialect cluster with a little over 2,100 users) is possibly the best described of them with its 1414 pp. grammar (Rice 1989).

[11] followed on the list by Ojibwa with Oji-Cree dialect (25,440 + 15,210), Inuktitut (40,320 speakers), Chipewyan (~Dene, 11,375), Micmac (9,000), Atikamekw (6,740; Algonquian, a variety of Cree, mentioned in the handbook only twice ("the school project that developed Wikipedia in the Atikamekw language", 799) but not in chapter ***40*** (cf. pp. 799, XV, and 1673); interestingly, Beland 1978:3 wrote that "the Atikamekw are less than three thousand and live in three villages"), Blackfoot (6,585), Slave (2,215). The 49-item list ends with <Tlingit (120)>.

[12] "The chapter discusses the genesis, status and terminology of the different languages that go by the name Michif" (951), cf. also "The language represented by the name [...] depends on the community in which it was spoken"; [...] the chapter "discuss[es] just one of these languages, [...] other languages also go by this name [...]" (953), despite the indicated minuscule but, on the other hand, growing population of speakers (according to *Statistics Canada*, the number of speakers for 2021 was 1,845, +57.7% from 2016; 13th place on the list, cf. fn 11).

[13] Probably omitted Stebbins 2003 deserved listing here.

43 "Wakashan Languages" by Tłatłakuł Patricia Rosborough (an adult learner of her late mother's language, Kwakwala", 1664) and Daisy Rosenblum (1013-52), family with seven languages of which some, thanks to Boas and Sapir, quite famous among linguists – like Nuu-chah-nulth (~Nootka; Ahousaht dialect mentioned in the handbook (53) but in the chapter only in "References" cf. 1050 under Nakayama 2003), Kwak'wala (~Kwakiutl), Heiltsuk (~Bella Bella), Nitinaht, Haisla – all seriously~critically endangered or no longer spoken;

44 Honoré Watanabe's "Salish" (1053-113), ~Salishan family embodying 23 languages, among them Nuxalk (~Bella Coola), Squamish, Shuswap, Coeur d'Alene (~Snchitsu'umshtsn), Comox (with Sliammon[14]), "most of them with further dialectal divisions" and "known for their phonetic and phonological complexity [...] and [...] rich morphology" (1053), most of them critically endangered or (Pentlatch, Tillamook, Twana~Tuwaduq~Skokomish, Quinault) extinct;

45 "Chinookan family, with special reference to Kiksht and notes on Chinuk Wawa" by Philip T. Duncan, Valerie (Lamxayat) Switzler, and Henry B. Zenk (1115-38); the last fluent speaker of Kiksht, the last Chinook -lect spoken, is said to pass away on July 11, 2012; Chinuk Wawa (~Wawa~Lelang) as an entity is also known to linguists but – under a different name: Chinook Jargon, mainly as an example of pidgin[15];

46 Joana Jansen's "Sahaptian" (1139-67), family with two languages: Sahaptin (~Ichishkiin, with two dialects and at least 13 further subdivisions, severely endangered) and Nez Percé (~Nimipuutimt, with two dialects, critically endangered, allegedly 20 speakers in 2007, but famous for its impressive over 1300pp. dictionary, with twenty unique photographs, by Haruo Aoki[16]);

47 "Karuk" (~Karok~Araráhih) by Andrew Garrett, Susan Gehr, Erik Hans Maier, Line Mikkelsen, Crystal Richardson, and Clare S. Sandy (1169-200), an isolate considered seriously endangered, yet we read that "in 2020, there are only a handful [allegedly 12 in 2007] of elder first-language speakers [b]ut there are fluent younger speakers who did not grow up fluent; and it is important to add that they and many others did grow up with the language around them. There has never been a time when Karuk was absent from every home [...], language classes are taught in [...] schools" and "community classes are offered [...]" (1194-5);

48 "Wá·šiw" by M. Ryan Bochnak, Emily A. Hanink, and Alan Chi Lun Yu (1201-21), better known in literature as Washo (also throughout the handbook)~Washoe and treated as an isolate but a number of other affinity suggestions emerged; 20 elderly native users quoted for 2008, current "revitalization efforts" reported;

49 Eugene Buckley's "Pomoan" (1223-46), family of seven languages, in literature practically all labeled <Pomo> with a, usually toponymic, determiner (Southeastern (~Clear/Lower Lake Pomo), Eastern, Northerastern, Northern (~Coyote Valley~ Little Noyo River), Central, Southwestern (~Kashiya), Southern (~West Creek~Dry Creek), etc.), five of them

[14] A 618-page grammar by Watanabe (2003) is worth mentioning here.

[15] The authors consider the "two languages [as being] actively spoken today" (1135, cf. also 1116f. with fn 2).

[16] Aoki wrote (1994: ix): „Today there are more than two thousand members of the Nez Perce tribe, but [...] the number of speakers of the Nez Perce language is not easy to estimate [...]. There are still hundreds of people who can count up to ten, but only scores can tell traditional folktales using classical vocabulary".

extinct, "most" considered "dormant", with "increasing interest in language revival and revitalization; [...] At this writing, there are a few native speakers of Central, and perhaps a dozen speakers of Kashiya" (1224, cf. also 1242-3);

50 Carmen Dagostino's "California languages: Isolates and other languages" (1247-74); according to "Tab. 1: California languages covered in this chapter" (1251-2), 37 variants of 15 head units (five isolates, seven families, three languages) have been covered, including e.g. Yokuts, Klamath, Modoc, Achumawi, Miwo, Wintu, Yana, Maidu, Wappo, classified also as either "dormant", "obsolescent", or "awakening" (in four cases combining two of these features)

51 Timothy P. Henry-Rodriguez's "Chumashan" (1275-302), family of six attested languages distinguished by their Hispanic glottonym endings (Obispeño, Cruzeño, Barbareño, Purisimeño, Samala (Ineseño), and Ventureño), all extinct: "By the mid-20th century, all Chumashan languages had lost their native speakers" (1275); one more remark seems worth quoting: "the list of completed dictionaries and grammars of Chumashan languages is short" (*ib.*);

52 Amy Miller's "Yuman" (1303-32), family with some ten ("recognized by U.S. linguists" and mostly "relatively well documented", 1303-4) to 22 and even 25 languages, including Kiliwa, Paipai, Hualapai~Walapai, Havasupai, Yawapai, Mohave, Maricopa, Cocopah~Kwapa), Cucapa~Kuapá, Kwatsáan, I(i)pai~ Diegueño, Northeastern Kumeyaay~ Diegueño, San José de la Zorra; "many [...] caught in a cycle of non-recognition and neglect: lacking recognition, they have not been systematically documented" (*ib.*), "endangered, in most cases severely or critically so" (1305), on the other hand, "many Yuman language communities have developed practical orthographies [, e]ach unique, reflecting the phonemic system of the language it represents and the preferences of its speakers" (1308); "other readings of interest" than "references" added (Langdon 1976 seems overlooked);

53 "Uto-Aztecan" by Eric Elliott and David Leedom Shaul (1333-59), family of about 30 languages, many of them known to linguists by their names, starting the list perhaps with the time/tenseless Hopi[17], also Paiute, Shoshone, Comanche, Cahuilla, Luiseño, Tubatulabal, Pima, Tepehua(n), Yaqui~Yoeme, Mayo, Huichol, Nahua(tl); individual languages as well as the entire family have been subject to intensive research to the extent that "(m)ajor works after 2000" turned out to be "too numerous to list" (1335); this writer would recommend to general linguists familiarization with subchapter 53.6 on "Vitality of Uto-Aztecan languages" (1348-56);

54 Logan Sutton's "Kiowa-Tanoan" (1361-406), family of seven or eight languages spoken in 13 or 14 communities" (1361, both listed on p. 1362): Kiowa, Towa, two Tewa languages, three Tiwa languages, and extinct Piro (one more Tiwa?); very informative and well prepared fragments indicating or recommending existing literature and on current "language situations" (1363-6, an impressive list of references 1391-406);

[17] Regretably, in this case the authors decided *against* adding "other readings of interest" listing such titles as e.g. Malotki 1983 (cf. the mottos opening the 700pp. volume), Karttunen 1983, Saxton et al. 1983, Robinson & Armagost 1990 or... Shaul 1999 and 2002.

55 "Caddoan" by L. Sutton and Armik Mirzayan (1407-46), "family of five documented languages": Caddo, Arikara, Pawnee, Wichita, and Kitsai (probably all dead or on the verge of extinction but "all of the contemporary Caddoan communities [save Kitsai] are [said to be] interested and engaged in language revitalization efforts" (1408-9); abundant recorded language data examples illustrating discussed issues is worth special appreciation;

56 A. Mirzayan's "Sketch of the Siouan Language Family" (1447-518, perhaps the most extensive sketch and chapter[18]), Family consisting "of *a number* of languages[19] – each with several dialects – spoken by at least 25 Indigenous Nations of North America in a broad area [...]" (1447, *italics* afm), including (Eastern) Dakota~Santee, Lakota (~Teton Sioux), Crow, Hidatsa, Assiniboine~Nakota, Hochunk~Winnebago, Omaha-Ponca, Osage, extinct Biloxi, Mandan, Tutelo, Saponi, and Ofo, and distantly related Catawba and Woccon; preservation statistics are contradictory and confusing, apart from <extinct>, several are "currently sleeping" or dormant, several seriously endangered, substantial population (4,160 for Crow in 2015, 2,100 for Lakota in a 2016 source[20]) seems exceptional;

57 Daniel W. Hieber's "Chitimacha" (1519-43), and

58 "Tunica" by Judith M. Maxwell and Patricia Anderson (1545-75) are sketches of two dead but sufficiently documented language isolates with revitalization attempts;

59 Jack B. Martin's "Muskogean" (1577-99), family of seven languages, some with names known to wider circles of linguists not specializing in Amerindian linguistics: Choctaw (9,600 users in 2015 census), Koasati, Alabama, Chickasaw, Seminole (Muskogee~Creek and Mikasuki~Hitchiti), and extinct Apalachee (the author adds here trade Mobilian Jargon, also extinct);

60 "Iroquoian" by M. Mithun and Ryan DeCaire (1601-25), family of, depending on the source, between nine and sixteen etnolects with the status of independent languages, including Cherokee (2,100 speakers in 2019, famous for its own unique writing system), Mohawk (with the population of speakers between 1,140 and 3,875 in 2016 considered "threatened"), Seneca, Cayuga, Onondaga, Oneida – all considered "seriously endangered"[21], and Wenro(hronon), Conestoga~Susquehannock, Tuscarora, Nottoway~Meherin, Huron~Wyantot, and some more, like Erie, Scahentoarrhonon, Laurentian), classified as "extinct";

61 Raoul Zamponi's "Unclassified languages" (1627-47); "a guide to what is known about fifteen languages once spoken [...] that now exist only as fragmentary records which

[18] "In general [the editors wrote in the "Preface"], we have opted for a greater number of shorter chapters rather than fewer longer ones, with the goal of covering as many relevant topics as possible while striving for user friendliness, though we recognize that the chapters necesarily vary somewhat in their accessibility and interest to different audiences" (VI).

[19] depending on the source, from 10 to 20, and more... ("... twenty two or so different Siouan languages portrayed...", p. 1451).

[20] A representative sample of Lakota (and some Caddoan Pawnee) could be heard worldwide by millions of spectators (thousands of linguists included) of Kevin Costner's 1990 film *Dances with Wolves*. Although much is said about surge of interest and revitalization, many evident educational aids (like e.g. WarClaud or Karol dictionaries) resulting from these trends have, unfortunately, not been listed in the bibliography (no list of "other readings of interest").

[21] And it is a euphemism in this writer's opinion.

resist (a convincing) classification": northeastern Nansemond and Pamunkey, southeastern Akokisa, Bayogoula, Bidai, Calusa, Congaree, Cusabo, Guale, Sewee, Shoccoree-Eno, and Amotomanco, Aranama, Solano, and Tanpachoa of the Southern Plains" (1628).

The handbook ends with three indexes – of languages (1673-84, as a tool well harmonizing with the front matter "list of families, languages, and dialects" mentioned at the beginning of this text), of names (1685-94), and of subjects (1695-702).

Spiritus moventes of the handbook under scrutiny here "intended to provide broad coverage of topics of interest to linguists in general, and more specifically to community and academic scholars engaged in the study and *revitalization* of North American languages. Particular attention has been given to new ideas and recent trends in research, to features of the languages that are typologically unusual or unusually well developed in comparison with others outside of the area, and topics of special importance to communities" (V, *italics* afm. – revitalization is *leitmotiv* of the entire publication).

For this reviewer, it is obvious that the handbook will find its place in personal libraries of the academic scholars mentioned and equally obvious expecting that most of particularly important works on North American Indigenous languages quoted as sources and listed in the bibliographies throughout the handbook are already in these private collections or at least remain easily accessible in nearby university libraries[22]. For native community scholars, activists, teachers, etc., it undoubtedly can and will long serve as a tool, a source of inspiration, a guidebook in their efforts to retain their in most cases seriously or critically endangered tongues (as well as other components of cultural heritage).

For – globally not small – flocks of "linguists in general" who have never even planned any involvement in dealing with Amerindian languages but always wished to widen their professional knowledge and understanding of "<language> in general" also through interest in the enormous variety of languages and linguistic structures and their relation(s) with the ways of thinking of their users, for linguists who have no libraries with rich collections of literature (grammars, dictionaries, handbooks, text anthologies) on – very intriguing for them – indigenous languages of North America[23], the handbook can potentially, and then successfully, serve as a reference book (together perhaps with additional support from publications like Heizer 1978), an encyclopedia of the discipline (Amerindian linguistics) with guaranteed competence, source for amazing (at times shocking) illustrative examples adding life and color to lectures and writings, an info-book to reliably help to decide whether, say, a modestly printed Rath 1981, Sylestine et al. 1993, DeBlois 1996 and CLC 1974 are reliable dictionaries, whether one can learn the language from e.g. Goossen 1995[24], or simply to provide linguistic facts and phenomena, and hints for further study (hence underlining above the

[22] It is not the case of "linguists in general" active in other branches or areas of linguistic research in most countries outside USA.

[23] Every course in linguistics includes, as a must, the "linguistic relativity (~(Sapir)-Whorf hypothesis" and almost every participant of such a course must have heard about the "Hopi timelessness" but relatively few linguists got acquainted with, or just had a look, at texts like Malotki's mentioned as an example in fn **17**.

[24] Rath and Sylestine et al. are listed ("reliable" for a user of the handbook), the other three not found, probably not mentioned (thus, possibly "unreliable").

importance of listing recommended additional or other relevant literature "of interest"). Using the material as a whole is not specially difficult when one gets familiar ~ learns how to optimize it (e.g. simultaneously using the indexes, table of contents, and the "List of North American families, languages, and dialects" mentioned above), although problems can emerge with details.

To demonstrate it, this writer selected only one such "problem" and only one chapter – that on "Inuit-Yupik-Unangan" – not in order to particularly criticize it (there is nothing peculiar inclining to "criticize") but rather to reveal some problems he came across within a period much too short[25] to identify all shortcomings possible in every publication. The selection, however, was deliberate: it is the only chapter which takes into account (or at least mentions) -lects from beyond North America[26].

His first reaction inducing some more detailed search throughout the two volumes was the paragraph (869) mentioning MacLean's Inupiaq 2014 dictionary as "particularly impressive – arguably the most comprehensive dictionary of any language of the family". Indeed, with its 1036 pp., 15~19,000 entries, an English index, a grammatical outline, bibliography, and thirty-one ethnolinguistic appendices, it *deserves* the attributive *impressive*, and there is no doubt about it. But the preceding paragraph (*ib.*) discusses "Yupik languages" in the very same context (the caption of the subchapter (38.7) is "Language maintenance and revitalization" (868-70)), and Jacobson's 1984, and especially [2]2012 Central Alaskan Yupik two-vols. dictionary which seems, arguably, equally impressive with its 1247 pp. "and approximately 11,200 entries (and subentries) in the main section" (p. 10[27]) have not been mentioned in the text or listed in the "References" (872)[28]. Of course, Rubtsova's 1971 dictionary (much less impressive in size but in fact with its 19,000 entries on 580 pp.[29] equally imposing) also not, possibly because of the Russian metalanguage of the publication.

[25] It is pity that we are not accustomed to, and do not expect reviews written a few *years after* the publication of works like the handbook here described.

[26] Perhaps, one more reservation is not out of place: as a principle, this reviewer avoids looking in similar texts for petty mistakes (like misprints, insignificant fact inexactitudes or misinterpretations) or suggesting the authors' "omissions" (it is the author's inalienable prerogative to select, omit, widen or limit, add or reject anything while creating her~his text), unless such remarks seem prospectively functional (a reviewer is neither proofreader nor editor).

[27] Jacobson warns that "this figure should be used with caution in making comparisons with other Eskimo dictionaries. Numerical comparisons will be meaningful only if the other dictionaries have been compiled following the same criterion (or more generally put, the same spirit of inclusion, the same judgment of non-predictability) that has been followed in compiling this dictionary" (*ib.*). The 1984 765pp. edition contained approximately 6,500 entries and subentries (*ib.*).

[28] Painstaking, but quick and one-time, poring over the handbook revealed at least four references to Jacobson's dictionary, two to its 1984 edition, in chapters 4 (on "Segmental phonology", 90, 106) and 22 (on "Pluractionality and distributivity", 512, 525) and two in chapter 27 (on dynamics of language systems, 634, 635, 645) to [2]2012. The "Index of names" proved not helpful (cf. p. 1688) which means that, optimally, in such monumental editions *every case of appearance* of personal names (also in references and other logistic tools or systems to make the edition user-friendly) should mandatorily be documented in such an index.

[29] supplemented with a "short index of stem- and form-derivative suffixes" including the really genuine attraction for linguists with forms~words derived from *qimuhsi-*, cf. fn 6.

Probably few, if any, potential users of the handbook under scrutiny, seeing and knowing its title, would reach for it anticipating comprehensive information on minor languages of northeastern recesses of Asia. On the other hand, since the territory of the Eskimo-Aleut language family does spread beyond America, it would be difficult to leave the fact unmentioned. Surprisingly, the chapter on these languages provides in short quite a lot of information in this respect, this reviewer expected one or two longer sentences or, at most, a short paragraph, signaling it only – but including references to at least the most important sources related to glotto- or topolects quoted like *Sirenik*, *Naukan*, *Central Siberian Yupik*[30] (in form of e.g. "other relevant literature ~ readings of interest" appended to "References" in some chapters) would not be out of place. In this respect, the quarter of a century earlier Mithun 1999: 400-3 is much better as a source, although references to Russian works are not made (except for Rubtsova 1971). Recommended is Dorais ^{3}2017.

The author of the Eskimo-Aleut chapter quoted (848) and listed in his bibliography (870-3) de Reuse 1994 (in which many references to important Russian source publications can be found), so he could treat it also as kind of substitutional solution assuming that one interested in such literature would in need reach just for that monograph which, however, is naturally a bit outdated and its bibliography of "References" (*ib.*, 459-80) is far from easy to use[31]. Recommended is Dorais ^{3}2017[32].

Full of praise for P.T. Authors, Editors, and Publishers of the handbook surveyed and portrayed here, below we shall wind up this text with a short independent premium illustrated list of references to selected literature concerning Eskimo-Aleut tongues from the opposite coast of the Bering Strait, also with focus on revitalization efforts there ("other readings of interest"), dedicated to readers of the present review.

Examples of important results of Russian academic research: **Yupik** – dictionary (Rubtsova 1971); grammars (Menovshchikov 1962, 1967; *Sirenik* 1964); texts (*Chaplino* Rubtsova 1954, Menovshchikov 1988; *Naukan* Menovshchikov 1987); monographs (Menovshchikov *Naukan* 1975; *Imaqliq* 1980; Vakhtin *Chaplino* 1987; 1995; *Old Sirenik* 2000; *New Chaplino* Yemelyanova 1982); **Inuit~Inupiaq / Imaqliq** – (Menovshchikov 1980); **Aleut** – (a). dictionaries Bering Island: – (Menovshchikov 1977; Oshima 2003: 1-308, 321-48); monograph (Golovko & Vakhtin & Asinovskiy 2009); (b). Copper Island (Golovko & Vakhtin 1990). Examples of school education aids as revitalization effort results: (a) teacher's aids – Menovshchikov & Vakhtin 1983; Vakhtin & Yemelyanova 1988[33]; (b). dictionaries – Golovko 1994

[30] Neither 'Central' nor 'Siberian'.

[31] Russian language text (titles, etc.) is provided in Roman character transliteration (not necessarily optimally chosen) with neither translation nor explanation given and the unusual medley of arrangement of the bibliography ("abbreviations are keyed to references found in the text" (*ib.*, 459, good question *why?*) and interwoven with the alphabetically arranged bibliographical descriptions) rather irritates than helps. No help can be expected also from e.g. Miyaoka 2012 (used and listed by Compton) quoting e.g. (pp. 3 and 16) Menovshchikov 1959 and 1964 without, however, identifying them in his list of "References/sources" (cf. *ib*., 1589-90, 1600-1).

[32] Available for this reviewer has been only the 1990 edition.

[33] There are at least six booklets – methodical programs and recommendations for teachers concerning the organization of Eskimo language courses for kindergarten children 2-, 3-, and 5-years of age and for 1st-4th primary (Eskimo!) school grade pupils published by central (Ministry of Education) and local state administrative authorities between 1986 and 1989 in Magadan, Provideniya, and Anadyr – in this reviewer's possession.

(Bering Island Aleut)[34]; Menovshchikov 1954; 1988 (Chaplino Yupik); (c). handbooks – Aynana et al. 1989; Menovshchikov 1974. Cf. also Menovshchikov 1968.

Menovshchikov 1974 Aynana et al. 1989

Menovshchikov 1988 Golovko 1994

Primary school books for Chukotka (Siberian) Yupik and Bering Island Aleut (Russia)

[34] The compiler of the dictionary and main author of Golovko, Vakhtin & Asinovskiy 2009 listed (p. 13) six names of his principal informants – all ladies, two of them deceased; the very same names have been listed in the 2009(:5) book, with the remark that "unfortunately, most of them no longer are among the living". On March 7, 2021, Golovko informed that the last native speaker of Bering Aleut, Vera Terentyeva Timoshenko, aged 93, passed away in her native village of Nikolskoye on Bering Island.

References[35]

~ 'or', alternative~variant
// (an)other language variant provided *in* the source
[ ] translation, explanation, intrusion, additional information by this reviewer

ANLC – Fairbanks: University of Alaska, Alaska Native Language Center (~Archives).
Aoki, Haruo. 1994. *Nez Perce dictionary*. Berkeley – Los Angeles – London: University of California Press.
[Aynana, Lyudmila Ivanovna & Gleb Aleksandrovich Nakazik & Marina Ivanovna Sigunylik] Л[юдмила] И[вановна] Айнана, Г[леб] А[лександрович] Наказик, М[арина] И[вановна] Сигунылик. 1989. *Юпигыт улюат 2* // *Эскимосский язык. Учебник и книга для чтения для 2 го класса. 2'е издание, доработанное* [Eskimo language, handbook and reader for the second grade primary school, 2nd edition]. LP.
Basse, Bjarne & Kirsten Jensen (eds.). 1979. *Eskimo languages: Their present-day conditions* [collection of lectures given at the symposium] *Majority Language Influence on Eskimo Minority Languages* [at the Department of Greenlandic, University of Aarhus, Oct. 1978]. Aarhus: Arkona.
Beland, Jean Pierre. 1978. *Atikamekw morphology and lexicon*. Berkeley: University of California. (Doctoral dissertation.)
Boas, Franz. 1911. *Handbook of American Indian languages... Part 1*. 1922. *Part 2*. Washington: Smithsonian Institition, Bureau of American Ethnology.
Bok-Bennema, Reineke. 1991. *Case and agreement in Inuit.* Berlin – New York: Foris Publications.
Campbell, Lyle. 1997. *American Indian languages: The historical linguistics of Native America*. New York: Oxford University Press.
Campbell, Lyle and Marianne Mithun (eds.). 1979. *The languages of Native America: Historical and comparative assessment*. Austin: University of Texas Press.
CLC – Carrier Dictionary Committee (Francesca Antoine, Catherine Bird, Agnes Isaac, Nellie Prince, Sally Sam) aand Richard Walker & David B. Wilkinson (SIL) 1974. *Central Carrier Bilingual Dictionary*. Fort Saint James, British Columbia: CLC.
CWBP – *The Collected Works of Bronisław Piłsudski* , vols. 1. *Aborigines of Sakhalin*; 2. *Materials for the Study of the Ainu Language and Folklore (Cracow 1912)* (1998); 3. *Materials for the Study of the Ainu Language and Folklore 2* (2004). Berlin – New York: Mouton de Gruyter; 4. *Materials for the Study of Tungusic Languages and Folklore* (2011). Berlin/Boston: De Gruyter Mouton.
DeBlois, Albert D. 1996. *Micmac dictionary*. Hull, Quebec: Canadian Museum of Civilization.
de Reuse, Willem Joseph. 1994. *Siberian Yupik Eskimo. The language and its contacts with Chukchi*. Salt Lake City: University of Utah Press.
Dorais, Louis-Jacques. 1990. *Inuit Uqausiqatigiit* // *Inuit languages and dialects*. Iqaluit: Arctic College – Nunatta Campus. Second, revised edition 2003. ^{3}2017. *Inuit languages & dialects* // *Inuit Uqausiqatigiit. Updated Edition*. Iqaluit: Nunavut Arctic College Media.
Dürr, Michael & Egon Renner & Wolfgang Oleschinski (eds.). 1995. *Language and culture in Native North America: Studies in honor of Heinz-Jürgen Pinnow*. München – Newcastle: Lincom.
Edwards, Keri. 2009. *Dictionary of Tlingit*. Juneau: Sealaska Heritage Institute. Cf. Naish & Story 1963 and Story & Naish 1973.
ELPR – Suita, Osaka: Endangered Languages of the Pacific Rim [research project].
Geoghean, Richard Henry. 1944. *The Aleut language: The elements of Aleut grammar with a dictionary in two parts containing basic vocabularies of Aleut and English*. Washington, DC: United States Department of the Interior. Reprint 1964, Seattle, WA: Shorey Book Store.
Goddard, Ives (ed.). 1996. *Handbook of North American Indians*, vol. 17. *Languages*. Washington DC: Smithsonian Institution.
[Golovko, Yevgeniy Vasilyevich] Е[вгений] В[асильевич] Головко 1994. *Словарь алеутско-русский и русско--алеутский (беринговский диалект). Пособие для учащхся начальной школы* [Aleut-Russian-Aleut primary school dictionary, 4,000 words, 320pp.]. СПб.: Просвещение [StPetersburg: Prosveshcheniye Publishers].

[35] and "other relevant literature ~ "readings of interest" used (including abbreviations).

Golovko, Eugeni V. & Nikolai B. Vakhtin 1990. Aleut in contact: The CIA enigma. *Acta Linguistica Hafniensia* 22. 97-125.

[Golovko, Yevgeniy Vasilyevich & Nikolay. B. Vakhtin & Aleksandr Semyonovich Asinovskiy] Е[вгений] В[асильевич] Головко, Н[иколай] Б[орисович] Вахтин, А[лександр] С[еменович] Асиновский. 2009. *Язык командорских алеутов. Диалект острова Беринга*. [Commander Archipelago Bering Island dialect of Aleut]. Санкт-Петербург «Наука» [St.Petersburg: Nauka].

Goossen, Irvy W. 1995. *Diné Bizaad: Speak, read, write Navajo*. Flagstaff, Arizona: Salina Bookshelf.

Heizer, Robert F. (ed.). 1978. *Handbook of North American Indians*, vol. 8. *California*. Washington, D. C.: Smithsonian Institution.

Jacobson, Steven A. 1984. *Yup'ik Eskimo-dictionary* (with "English-to-Yup'ik Index" (693-757)); [2]2012. *Yup'ik EskimoDictionary*. *Second Edition.* vols. 1-2 (with "English-to-Yup'ik Index" (1013-247)). ANLC.

Jacobson, Steven A. (ed.) & Linda Womkon Badten & Vera Oovi Kaneshiro & Marie Oovi [2]1987. *A dictionary of the St. Lawrence Island/Siberian Yupik Eskimo language*. ANLC.

Karol, Joseph S. & Stephen L. Rozman (eds.). 1971, 1974 (first revision), 1997 (second revision[36]). *Everyday Lakota: An English-Sioux dictionary for beginners*. St.Francis, S.D.: Rosebud Educational Society.

Karttunen, Frances. 1983. *An analytical dictionary of Nahuatl*. Norman and London: University of Oklahoma.

Krauss, Michael E. 1980. *Alaska native languages: Past, present, and future*. ANLC.

Krauss, Michael E. (ed.). 1985. *Yupik Eskimo prosodic systems: Descriptive and comparative studies*. ANLC.

Langdon, Margaret. 1976. *Yuman texts* (Native American Texts Series). Chicago: University of Chicago Press.

Langdon, Margaret & Shirley Silver (eds.). 1976. *Hokan studies: Papers from the First Conference on Hokan Languages held in San Diego, California, April 23-25, 1970*. The Hague – Paris: Mouton.

Leer, Jeff. 1978, [2]1982. *A conversational dictionary of Kodiak Alutiiq*. ANLC.

LN – Leningrad: Nauka // Ленинград: Издательство "Наука".

LP – Leningrad: Prosveshcheniye // Ленинград: Издательство "Просвещение".

MacLean, Edna Ahgeak. 2014. *Iñupiatun Uqaluit Taniktun Sivuniŋit // Iñupiaq to English dictionary*. ANLC. Cf. also L.-J. Dorais's review in *Anthropological Linguistics* 56(3). 119-20.

Malotki, Ekkehart. 1983. *Hopi time: A linguistic analysis of the temporal concepts in the Hopi language*. Berlin – New York – Amsterdam: Mouton Publishers.

Mather, Elsie & Marie Meade & Osahito Miyaoka. 2002. *Survey of Yupik grammar revised*. ELPR.

[Menovshchikov, Georgiy Alekseyevich] Г[еоргий] А[лексеевич] Меновщиков. 1954. *Эскимосско-русский словарь* [with 4,700 words, a "suffix list" and a grammatical outline]. Ленинград: Государственное учебно-педагогическое издательство Министерства просвещения РСФСР, Ленинградское отделение [predecessor of LP].

Menovshchikov, G. A. 1959. Эскимосско-алеутские языки. Распространение и классификация [Eskimo-Aleut languages, geographic range and classification]. In Е. А. Бокарев и Ю. Д. Дешериев [Ye. A. Bokarev & Yu. D. Desheriyev] (eds.), *Младописьменные языки народов СССР* [languages of USSR with recently introduced writing systems], 300-17. Москва-Ленинград: Издательство Академии наук СССР [Moscow & Leningrad: USSR Academy of Sciences Publishing House].

Menovshchikov, G. A. 1964. *Язык сиреникских эскимосов. Фонетика, очерк морфологии,тексты и словарь* [Sireniki Yupik outline of phonetics and morphology (3-106), texts (107-75), and vocabulary (176-203(-16)]. ML & LN.

Menovshchikov, G. A. 1962. *Грамматика языка азиатских эскимосов. Часть первая. Фонетика, морфология именных частей речи* [Yupik grammar, part 1: phonetics and morphology of declinables]. 1967. *Грамматика языка азиатских эскимосов Часть вторая Глагол, причастие, наречия, служебные слова* [part 2: verb, participle, adverbs, auxiliary words; pp. 588 (300+ 288).]. LN.

Menovshchikov, G. A. 1968. Эскимосско-алеутская группа [Eskimo-Aleut group] (pp. 352-65), "Эскимосский язык" [Eskimo language] (366-85), and Алеутский язык [Aleut language] (386-407). In П. Я. Скорик (отв. ред.), *Языки народов СССР. Том пятый. Монгольские, тунгусо-маньчжурские и палеоазиатские языки* [P Ya Skorik et al. eds. "Languages of the USSR, vol. 5. Mongolic, Manchu-Tungusic, and Paleoasiatic languages]. LN.

[36] "This book includes 3,800 entries, 300 phrases, idiom drills, expressions of time, coinage, native birds and animals, and rules for forming Lakota sentences" (quoted from the title page).

Menovshchikov, G. A. 1974. *Учебник юпигыт улюн'истун аюк'ылг'и подготовительнымун классымун* [Eskimo language handbook for preparatory grade~course]. LP.

Menovshchikov, G. A. 1975. *Язык науканских эскимосов (фонетическое введение, очерк морфологии, тексты, словарь)* [Naukan Yupik: outline of phonetics and morphology (4-341), texts (342-78), dictionary (479-510)]. LN.

Menovshchikov, G. A. 1977. Алеутско-русский словарь [Aleut-Russian vocabulary]. *.Языки и топонимия* (Tomsk) 5. 137-98.

Menovshchikov, G. A. 1980. *Язык эскимосов Берингова пролива* [Bering Strait Eskimo – grammar (22-191), texts (192-217), vocabulary (217-325)]. LN.

Menovshchikov, G. A. 1987. *Материалы и исследования по языку и фольклору науканских эскимосов* [materials for the study of Naukan Yupik – texts with commentaries (243pp.)]. LN.

Menovshchikov, G. A. 1988. *Словарь эскимосско-русский и русско-эскимосский Пособие для учащхся начальной школы* [primary school Yupik-Russian-Yupik dictionary, 4,000 words]. LP.

Menovshchikov 1988. *Материалы и исследования по языку и фольклору чаплинских эскимосов* [materials for the study of Chaplino Yupik (236pp.)]. LN.

Menovshchikov, Г. A. & N. B. Vakhtin [cf below]. 1983. *Эскимоский язык* [Eskimo language, handbook for students of pedagogical colleges]. LP.

Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge: Cambridge University Press.

Miyaoka, Osahito. 2012. *A grammar of Central Alaskan Yupik (CAY)*. Berlin – Boston: De Gruyter Mouton.

MN – Moscow: Nauka // Москва: Издательство Наука.

Molina, Felipe S. et al. – see Shaul 1999.

Nagai, Kayo. 2001. *Mrs. Della Waghiyi's St. Lawrence Island Yupik texts with grammatical analysis by Kayo Nagai*. ELPR.

Naish, Constance & Gillian Story 1963. *English-Tlingit dictionary: Nouns* [100pp., classified, with notes on phonetics and spelling]. Fairbanks, Alaska: Summer Institute of Linguistics. Cf. Edwards 2009 and Story & Naish 1973.

Oshima, Minoru. 2003. *Материалы об алеутском языком* [*sic!*] *на острове Беринга: Словарь и тексты* // *Linguistic materials of Bering Island Aleut. Dictionary and Texts* metalanguages are Russian and English, 5 texts and twelve dialogues included, 309-20]. Suita, Osaka: ELPR.

Rath, John C. 1981. *A practical Heiltsuk-English dictionary with a grammatical introduction: Volume*[s] *One & Two* [over 760 pp.]. Ottawa: National Museums of Canada.

Rice, Keren. 1989. *A grammar of Slave* (Mouton Grammar Library 5). Berlin – New York: Mouton de Gruyter.

[Rubtsova, Yekaterina Semyonovna] Е[катерина] С[еменовна] Рубцова. 1954. *Материалы по языку и фольклору эскимосов (чаплинский диалект)* [materials for the study of Chaplino Yupik, texts with commentaries]. Москва – Ленинград: Издательство Академии наук СССР [Moscow – Leningrad: USSR Academy Press].

Rubtsova, Ye. S. 1971. *Эскимосско-русский словарь* [Yupik-Russian dictionary, with 19,000 words on 578 pp., and lists of suffixes and of derivatives of *qimuhsi-* pp. 593-644; cf. fns 6 and 29]. Москва: Советская энциклопедия [Moscow: Soviet Encyclopaedia Pubnlishers].

Saxton, Dean & Lucille Saxton & Susie Enos 1983. *Dictionary Tohono O'odham/Pima to English English to Tohono O'odham/Pima*. 2 edn. / Revised and Expanded. Tucson: University of Arizona Press.

Shaul, David Leedom[37]. 1999. *Hippocrene standard dictionary Yoeme-English English-Yoeme with a comprehensive grammar of Yoeme language*. New York: Hippocrene Books.

Shaul, David Leedom. 2002. *Hopi traditional literature*. Albuquerque: University of New Mexico Press.

Statistics Canada / Statistique Canada (release date March 29) 2023. *Indigenous languages in Canada, 2021*. Infographics.

Stebbins, Tonya. 2003. *Fighting language endangerment: Community directed research on Sm'algyax (Coast Tsimshian)*. ELPR.

[37] Thus on book covers; the title page indicates Felipe S. Molina, Hermina Valenzuela, and David Leedom Shaul.

Sylestine, Cora & Heather K. Hardy & Timothy Montler & contributors Ivey Battise, Dorcas Bullock, Vincent Celestine, Wanda Poncho & James Sylestine) with the assistance of Jack Martin 1993. *Dictionary of the Alabama language*. Austin: University of Texas Press.

Story, Gillian L. and Constance M. Naish. 1973. *Tlingit verb dictionary*[38] [English-Tlingit-English, with "Grammar sketch", focusing on verb and verb phrase, appended]. ANLC. Cf. Edwards 2009 and Naish & Story 1963.

[Vakhtin, Nikolay Borisovich] Н[иколай]. Б[орисович] Вахтин. 1987. *Синтаксис простого предложения эскимосского языка* [Chaplino Yupik simple sentence syntax]. LN.

[Vakhtin, N. B.]. 1995. *Синтаксис языка азиатских эскимосов* [Chaplino Yupik syntax]. St.-Peterburg: Издательство Европейского Дома.

[Vakhtin, N. B.]. 2000. *Язык сиреникских эскимосов: Тексты и словарные материалы* // *The Old Sirenik Language: Texts, Grammatical Notes, Lexicon*. München: Lincom Europa.

[Vakhtin, N. B. & N. Y. Yemelyanova (cf. below)]. 1988. *Практикум по лексике эскимосского языка: Учебное пособие для педагогических училищ* [practical aid for teaching and learning New Chaplino Yupik lexicon]. LP.

WarCloud, Paul. 1971. *Sioux Indian dictionary: Over 4,000 words: Pronunciation at a glance*. Pierre, SD: State Publishing Company. [4]1989. *Dakotah Sioux Indian dictionary: Over 4,000 words: Pronunciation at a glance*. Sisseton, SD: Tekakwitha Fine Arts Center.

Watanabe, Honoré. 2003. *A morphological description of Sliammon, Mainland Comox Salish, with a sketch of syntax*. ELPR.

Wistrand-Robinson, Lila & James Armagost. 1990. *Comanche dictionary and grammar*. The Summer Institute of linguistics and The University of Texas at Arlington.

[Yemelyanova, Nina Mikhaylovna] Н[ина] М[ихайловна] Емельянова. 1982. *Классы глаголов в эскимосском языке* [verb classes in New Chaplino Yupik]. LN.

[38] A model potential question from inquisitive student audience: *don't they have nouns?*