LINGUA POSNANIENSIS
LXVI
2024 (1)

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

LXVI (1)

POZNAŃ 2024

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

POZNAŃSKIE TOWARZYSTWO PRZYJACIÓŁ NAUK
WYDZIAŁ FILOLOGICZNO-FILOZOFICZNY
KOMISJA JĘZYKOZNAWCZA
we współpracy
z UNIWERSYTETEM IM. ADAMA MICKIEWICZA W POZNANIU

LINGUA
POSNANIENSIS

CZASOPISMO POŚWIĘCONE JĘZYKOZNAWSTWU
PORÓWNAWCZEMU I OGÓLNEMU

LXVI (1)

POZNAŃ 2024

WYDAWNICTWO POZNAŃSKIEGO TOWARZYSTWA PRZYJACIÓŁ NAUK

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

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.

ISSN 0079-4740
eISSN 2083-6090

Wersja drukowana czasopisma jest wersją pierwotną.

TABLE OF CONTENTS

DOI: 10.14746/linpo.2024.66.1.1

The use of discourse markers in conversation: The case of "myślę" – "I think" in Polish Career Clean Coaching

Konrad Juszczyk[1] & Barbara Konat[2]

[1] Adam Mickiewicz University in Poznań
juszczyk@amu.edu.pl | ORCID: 0000-0002-4769-0155
[2] Adam Mickiewicz University in Poznań
bkonat@amu.edu.pl | ORCID: 0000-0003-2370-4636

Abstract

People find it difficult to talk about abstract concepts such as their future career plans. The discourse marker "I think" is used to express uncertainty and soften the force of a proposition. The objective of this paper is to demonstrate how specific context and bounded topic of conversation influence the use of the discourse marker "I think" in Polish (*myślę*) as a structuring discourse marker. A total of 23 participants were recorded during Career Clean Coaching sessions and the usage of "I think" in the resulting corpus was compared to other Polish conversational corpora. The qualitative study of the semantics of "I think" indicates that this discourse marker serves a hedging function. The quantitative analysis of the frequency of the discourse marker "I think" in Polish conversational corpora and in the Career Clean Coaching corpus reveals significant differences. The results indicate that pragmatic factors, namely the topic, context and protocol of the conversation, exert a strong influence on the usage of "I think".

Keywords: discourse markers, conversations, corpus study, clean coaching, pragmatics

1. Introduction

This paper presents the Career Clean Coaching corpus and compares it with two other Polish corpora: goal-oriented dialogue and a sample of natural, everyday conversations. The aim is to explore how the context of conversations influences the frequency of "I think" as a marker of uncertainty. To achieve this, we provide a detailed analysis of the Career Clean Coaching sessions and corpus, as well as a semantic distinction of the five basic senses of "I think" in Polish, which were annotated in the Career Clean Coaching corpus. Our initial study of coaching sessions has suggested that "I think" is

a keyword with high frequency, among other cognitive verbs (Konat & Juszczyk 2015). This initial finding motivates the analysis of "I think" as a discourse marker in the current paper.

This paper is structured as follows: Section 2 presents related work concerning the role and functions of discourse markers in English and Polish in different communicative contexts. Section 3 provides a detailed description of the data used in the study. In this section, the features of coaching as conversation are presented and the Clean Coaching protocol is introduced, along with the introduction of the Career Clean Coaching (CCC) corpus as a material. The fourth section presents the results of Studies 1-3.

Firstly, it is predicted that the participants (coachees) of Career Clean Coaching sessions will use the discourse marker "I think" mainly to soften the strength of their claims. The first hypothesis is that among all possible senses of "I think" in Polish, the most frequent will be the one associated with the function of a hedge. This is illustrated in our Study 1: Five senses of "I think" in Career Clean Coaching.

Secondly, we posit that the frequency of usage of "I think" is not solely a characteristic of an individual speaker but is also shaped by the form of conversation. Our second hypothesis is that the specific protocol, as well as abstract topics of Career Clean Coaching sessions, will result in a high frequency of "I think". To test this, we compared the Career Clean Coaching corpus with the corpora of other Polish conversations in Study 2: The frequency of "I think" in Career Clean Coaching.

Thirdly, we posit that although the usage of "I think" varies between speakers, the context, topic and specific protocol exert a stronger influence. This is explored in Study 3: Speaker differences in Career Clean Coaching, where inter-speaker variability is examined.

2. Related work

The research presented here is informed by theories of cognitive linguistics, conversation analysis and studies on discourse markers. The functions of discourse markers may vary depending on the communicative context in which they are used (Tay 2013: 164). They may indicate solidarity, rapport, and in-group membership (Cutting 2000), contribute to the perceived naturalness of conversation (McCarthy 1998), and convey a general sense of indirectness, often interpreted as being more polite (or less intrusive) in conversation (Brown & Levinson 1987). Tay's analyses of psychotherapy extracts suggest that discourse markers may be directly motivated by metaphor. In some cases, they may signal upcoming metaphorical expressions. Although Tay mentions that his analysis is not validated on large corpora, he also gives reasons why the use of discourse markers and metaphors may be interrelated. Metaphors are employed in conversation to convey evaluations, attitudes, values, perspectives, and beliefs (Cameron & Maslen 2010). Additionally, they are utilized when individuals "struggle to find words to capture difficult to describe sensations" (McMullen & Conway 1996: 252). Discourse markers such as "you know" or utterance-final "right" occur at significant information state transitions, where the objective of the speaker is to reduce an initial asymmetry with regard to some body

of knowledge held by the speaker and hearer (Tay 2013: 161). In this paper, we argue that "I think" in coaching sessions serves similar functions.

The data-driven approach, similar to conversation analysis of psychotherapeutic sessions (Ferrara 1994; Pawelczyk 2011; Tay 2013), is employed in this research. The analysis of coaching sessions as dialogues is framed in conversation analysis, as its scope is a professional discourse, and it relies on recorded interactions (sessions) of naturally occurring speech. The discourse marker "you know" has been identified as a facilitator of intimacy between clients and psychotherapists. This is evidenced by the observation that it "mitigates the threat posed to the client in verbalizing highly intimate information" (Pawelczyk 2011: 154). The marker is often used to precede a potentially threatening or traumatic thought or idea that is about to be revealed by the client. (Pawelczyk 2011: 151). This is why discourse markers such as "you know" often signal self-disclosure of the patient in psychotherapy.

Our study is also closely related to work on discourse markers of uncertainty in English and Polish. In this area, studies of "I think" have shown its versatile functions. In her study of rape trial discourse, Ponterotto (2014) included the phrase "I think" in the category of hedges, which are used by speakers when reporting difficult and traumatic experiences. Mulllan (2010) posits that the extensive use of "I think" is a strategy for dealing with uncertainty because it mitigates or softens the force of a proposition.

> The reasons for downtoning a proposition can include being unsure as to the truth or reliability of a statement, wishing to claim or disclaim responsibility for a proposition, and mitigating or softening a statement out of politeness or consideration for the addressee (Mullan 2010: 59).

In English conversations, "I think" is treated as an indicator of taking an epistemic stance (Kärkkäinen 2003). Spoken corpora have been employed to investigate the frequency of the use of "I think" as an indicator of uncertainty (Brezina 2013). The use of cognition verbs in the first-person singular may be regarded as a distinctive feature of conversational interactions. Helasvuo (2014) observed that in Finnish conversations, the verb "ajatella" – "to think" – appeared among the five most frequent cognition verbs, with a total of 132 instances. Among the instances, as much as 90 (68%) were used in the first-person singular form. Consequently, this paper focuses on the exact word form "myślę" – "I think", in which speakers are referring to their own cognitive states.

One of the cases where the pragmatic factor – a specific context of usage – influences the function of "I think" is the use of English either as a first language or second language. Baumgarten & House (2010) analysed the functional profiles of "I think" and "I don't know" in conversations between native speakers and in English as a lingua franca interaction. Biber (2012) provided quantitative evidence for the influence of the discourse register on the frequency of lexemes and structure, demonstrating how conversations differ from information writing. In this paper, we conduct a comparison between a balanced corpus, which is mostly written, and our coaching session corpus, confirming Biber's observation. We also provide evidence that pragmatic factors of conversation, such as setting, topic, and speakers, can strongly influence the lexical choices of the speaker.

Polish is a null-subject language. Some researchers (see, for example, Posio 2014) have argued that the presence or absence of an overt subject in the conversational use of "I think" should be considered in analyses. However, in the current study, we did not analyse the presence of an overt subject, as it was very rare (only 10 instances), which did not allow for generalisations. The position of "I think" in the phrase, either initial (Kaltenböck 2009) or final (Sato 2017), has been studied as a meaningful factor. However, this aspect is not considered in the current study.

The study presented here demonstrates that the usage of "I think" in natural, spoken communication, where the owner of the meaning of "I think" is the actual speaker, differs from that in official or written text. The method employed in our study is corpus analysis of language data, which is sometimes referred to as "a combination of interactional sociolinguistics and corpus linguistics" (Baker 2010). Finally, this paper aims to contribute to the study of the pragmatics of Polish spoken language. A study conducted on a sample of 100 articles from the Journal of Pragmatics revealed that only one article in the sample addressed Polish as a subject of study. Polish was identified as one of 20 languages with a single publication, whereas English was studied in as many as 50% of the articles (Egbert et al. 2016).

A number of studies have examined the use of the discourse marker "I think" in different languages and contexts. A multimodal analysis of the discourse marker "I think" in Brazilian Portuguese was conducted on a sample of audio-video recorded interviews (Freitag et al. 2021). The use of "I think" and similar phrases was examined in task-based group conversations with native and non-native Swedish speakers (Tolvanen 2024). The translation of the phrase "I think" from English to Lithuanian and Hebrew was examined using the alignment model of phrase-based statistical machine translation, with manual treatment of the data (Oleskeviciene & Liebeskind 2021). In conclusion, these studies demonstrate the multifaceted and context-dependent nature of the discourse marker "I think".

3. Material – Coaching Sessions

3.1. Coaching session

A coaching session is a conversation between a coach and a coachee, during which the coach employs a specific protocol and mirroring technique to encourage the coachee to self-disclose personal and professional experiences. Coaching is provided to employees in corporations, entrepreneurs, members of non-governmental organisations, and students in career advising offices. Individuals engage with a coach when they require assistance in developing self-esteem and self-confidence, as well as in enhancing their soft skills, including interpersonal communication, public speaking, work performance and work-life balance, time and team management, career change and other areas. The primary reason for selecting coaching is the perception that one's performance and success in life do not align with the expectations of one's supervisor or oneself.

In contrast to other forms of support, such as counselling, psychotherapy, mentoring or training, the coach does not typically provide advice to the coachee on how to change their attitude or performance. The individual undergoing coaching is referred to as a "client" rather than a "patient". The coach is not a therapist, as the coach does not diagnose the client and does not provide them with solutions, advice, or treatment. Instead, the coach guides the coachee through the process of change, with the objective of instilling self-steering. The role of the coach is to provide support to the coachee in understanding the nature of the problem, identifying the desired outcome and the necessary resources, and facilitating the coachee's willingness to change. The coachee is encouraged to provide a summary of their experience and to interpret the content of the session. This process enables the coachee to gain self-steering and independence in decision-making and action-taking. During the coaching session, the coachee is encouraged to explain their situation, identify possibilities, make decisions and take actions independently. Furthermore, coaching differs from psychotherapy in that the coach is focused on the coachee's future, rather than their past experience. The coach also draws the coachee's attention to their resources and helps them to find a solution that is suitable for their situation (Thorpe & Clifford 2003; Pawelczyk 2011; Sullivan & Rees 2008).

3.2. Coaching sessions model

In this section, we compare coaching with psychotherapy and everyday conversation. Since coaching is a form of providing support in conversation, we assume that it is similar to consultation or psychotherapy. To contrast coaching discourse with other kinds of conversation, we use Ferrara's model for differentiating psychotherapy from ordinary conversation (Ferrara 1994). Ferrara considers regular conversation to be the unmarked form of discourse (Ferrara 1994). In regular conversation, both speakers contribute equally to the content and details. The way in which speakers exchange information depends on the specific needs of the individual and the type of information being shared. For example, friends or family members meet to share their experiences and to spend time together, so the aim, form and topic of conversation are not explicitly stated, and the form is less regulated. According to Ferrara's model, there are seven dimensions which distinguish psychotherapy from conversation:

1. **Parity**
2. **Reciprocality**
3. **Routine recurrence**
4. **Bounded time**
5. **Restricted topic**
6. **Remuneration**
7. **Regulatory responsibility**

It can be argued that a coaching session differs from a conversation in the same way as a psychotherapy session does (see Table 1). All seven features of psychotherapy discourse are present in coaching discourse. We propose the application of Ferrara's model

to coaching, which allows us to present how a coaching session is similar to a psychotherapy session.

1. Parity: on this dimension, a coaching session is uneven, with a lower parity in comparison to everyday conversation. The client (coachee) agrees to be supported by the coach and allows the coach to conduct the conversation according to a certain protocol.

2. Reciprocality: This is absent in coaching, as only the coach inquiries about the coachee's experiences and attitudes, rather than the converse.

3. Routine recurrence: Coaching sessions constitute a fixed element of the coaching process, occurring at regular intervals and for a pre-determined duration. These meetings are a recurring routine for both parties.

4. Bounded time: Coaching sessions typically span a duration of between 30 and 90 minutes.

5. Restricted topic: The **topic** of the coaching session may be selected by the client (coachee) or proposed by the sponsor (company funding the coaching). Depending on the coaching type, the topic may relate to management skills, decision-making and task delegation in business coaching or work-life balance, mental health and self-fulfilment in life coaching, for example.

6. Remuneration: Coaching differs from conversation in that it is a professional service provided by trained specialists. Coaches are remunerated either by the client or by a sponsor.

7. Regulatory responsibility: The coach is responsible for the form of discourse development since the coach is the one who opens the session, forms questions for the coachee and terminates the session. However, the coach is focused on the coachee's career and experience, and therefore the specific content of the session is dependent on the coachee. It is not permitted for the coach to introduce any details of their personal life or professional career.

In order to describe Career Clean Coaching sessions, we propose the development of a new model based on Ferrara's existing model. Our new model incorporates three additional dimensions: self-disclosure, mirroring, and protocol. The last two dimensions have different values in Clean Coaching and psychotherapy, thus constituting the distinctive features of Clean Coaching.

8. Self-disclosure: in psychotherapy, coaching or everyday conversations, self-disclosure is understood as "the process of deliberately revealing information about oneself that is significant and that would not normally be known by others" (Adler & Towne 1996: 358). However, as Pawelczyk observes, "in psychotherapy, unlike in other social contexts, it is the client who is granted extensive clinical and conversational space to disclose" (Pawelczyk 2011: 123). Similarly, the coaching session represents another social context in which the coach provides the client with a conversational space to disclose and freely discuss their experiences and expectations in life. In regular conversation, self-disclosure can be mutual and equal, with both parties sharing their experiences and exchanging news about their lives.

9. Mirroring is an important aspect of building rapport and occurs naturally and spontaneously in conversation (Chartrand & Bargh 1999). Typical mirroring is manifested through similar body posture, gestures, intonation, and phrases (Kipp 2010). Some

psychotherapists and coaches are trained in the practice of mirroring and employ it deliberately (Thorpe & Clifford 2003). Psychotherapists frequently repeat their patients' words with the intention of fostering their patients' self-expression and indirectly requesting further elaboration (Pawelczyk 2011: 163). Similarly to psychotherapists, coaches combine exact repetitions or paraphrases of their clients' words with questions to draw attention to salient issues of their clients' experiences. This encourages clients to share their experiences in the process of self-disclosure (Thorpe & Clifford 2003). Such verbal mirroring is also referred to as back-channelling or paraphrasing (Pawelczyk 2011).

10. Protocol: a set of rules and speech acts that are typically employed by dialogue participants. However, in contrast to regular conversation, where the protocol is not explicitly known, the protocol used in psychotherapy or coaching is maintained by the psychotherapist or coach, respectively. The specific protocol governing the conduct of psychotherapy or coaching sessions varies according to the therapy type and determines the general aim and type of support provided. The specific protocol is explicitly known to the coach but not to the coachee. The detailed protocol of coaching sessions on which our study was based is described in the next section (Clean Coaching Protocol).

Table 1: Comparison between conversation, psychotherapy session and Clean Coaching

Dimensions	Conversation	Psychotherapy	Clean Coaching
1. PARITY	high (dialogue partners are equal)	low (in favour of psychotherapist)	low (in favour of coach)
2. RECIPROCALITY	high (dialogue partners are equal)	suspended	suspended
3. ROUTINE RECURRENCE	no	yes	yes
4. BOUNDED TIME	no	yes	yes
5. RESTRICTED TOPIC	no	yes	yes
6. REMUNERATION	no	yes	yes
7. REGULATORY RESPONSIBILITY	dialogue partners are equal	psychotherapist	coach
8. SELF-DISCLOSURE	mutual and equal	patient's	coachee's
9. MIRRORING	natural and spontaneous	performed by trained psychotherapist in form of paraphrase	performed by trained coach in form of parrot-phrasing (exact wording)
10. PROTOCOL	does not need to be stated and known explicitly	depends on the type of therapy	Clean Language

Table 1 summarizes how the coaching session is different from regular conversation, but similar to the one between psychotherapist and patient in psychotherapy (Juszczyk 2017).

3.3. The Clean Coaching Protocol

Clean Coaching is a type of coaching that employs a special set of questions and techniques, collectively known as Clean Language, which constitute the protocol for the session. Clean Language questions were developed by David Grove for psychotherapy in the late 1980s and subsequently established as a system by Lawley and Tompkins (2000). The protocol is currently employed worldwide in coaching, consultancy, conflict resolution, interviewing or interrogation, and business management and education (Sullivan & Rees 2008). In the current study, the Polish adaptation of Clean Language is utilised (Pieśkiewicz & Kołodkiewicz 2011). In this method, the coach selects questions from a limited set that facilitate the coachee in reflecting on their experience and developing metaphors of their experience. Metaphors are believed to facilitate discourse on complex and abstract topics, such as the future, development, or job-seeking (Sullivan & Rees 2008).

In Clean Language, questions are designed to be "clean", which entails that the coach does not propose any metaphorical expression but repeats the coachee's words without modification. This method of questioning is referred to as parrot-phrasing, in contrast to paraphrasing, where the coach modifies and reinterprets the coachee's words. Parrot-phrasing results in numerous repetitions of the coachee's phrases during the session. However, this approach simultaneously ensures that the coachee remains consistent and coherent in their responses, focusing on a specific topic.

The prescriptive protocol presented in Table 2 is reconstructed on the Clean Language handbook for practitioners (Sullivan & Rees 2008). The session commences with the Opening Question (also referred to as the Intention Question), which is typically of the following form: "What would you like to have happen?" or "When you think about your career, this is like what?". Subsequently, the coach repeats the coachee's words and requests further information using development questions, such as "Is there anything else about X?" or "What kind of X is that X?" where X represents the client's previous statements. Once the coachee has developed a metaphor for their career, the next stage of the session involves the use of so-called attributes and location questions. The coach may inquire as to whether the concept in question has a shape or size, or alternatively, where it is located. These questions are designed to assist the coachee in identifying connections and associations between the concepts. As more metaphors emerge in the coachee's words, the coach may pose the Relationships Question, for instance, "Is there a relationship between X and Y?". At this stage, the coach may also introduce Sequence and Source Questions, which are of the following form: "What occurs immediately preceding X?" or "What occurs next after X?" to conclude the session. The final stage of the session may include the questions, "What is the first thing you know now?" and "What is the second thing you know now?". This series of questions continues until the coachee responds to the question about the sixth thing the coachee knows as a result of the session. The question typically comprises the following elements: a conjunction (and), a pronoun (that, this, these, those), the coachee's words, and a clean question, followed by the coachee's words again. An example of a complete question is "And that eagle: what kind of eagle is that eagle?" or "And that rollercoaster: is there anything else about

that rollercoaster going fast?" or "And that surprise: does that surprise have a shape or a size?". This general protocol of Clean Coaching was used to obtain the Career Clean Coaching corpus for the current study.

Table 2: Overview of the protocol of Clean Coaching session

Type of question	General form of the question
Intention Question (Open)	"What would you like to have happen?" "When you think about your career this is like what?"
Development Questions	"Is there anything else about X?" "What kind of X is that X?"
Attribute and Location Questions	"Does X have a shape or size?" "Where is X?"
Relationships Questions	"Is there a relationship between X and Y?"
Sequence and Source Questions	"What happens just before X?" "When X what happens next?"
Knowing Questions (Terminate)	"What is the first (second, third etc.) thing you know now?"

It can be argued that these dimensions of coaching sessions set them apart from everyday conversations. These differences in the pragmatic set up of the interaction influence the lexical choices of speakers. In particular, the frequency of "I think" as a discourse marker is worthy of further investigation.

3.4. Clean Career Coaching corpus

The Clean Career Coaching corpus comprises 23 dialogues concerning the future careers of participants. These dialogues were recorded during free coaching sessions on professional development and career planning. The sessions were advertised via university career offices and a project website. The coaching sessions were conducted according to the model presented in Section 3.2 and the set of Clean Questions presented in Section 3.3. The participants were between the ages of 25 and 35 and were currently enrolled in or had recently graduated from a Master's or PhD program. This demographic may be expected to have concerns about their careers, as some had recently begun their professional lives, while others were still engaged in the job search process. Each subject participated in two coaching sessions with one of two hired coaches, who assisted them in defining career objectives and making decisions about their future careers. The coaching sessions were recorded with multiple cameras and microphones. None of the researchers were present in the recording room during the actual sessions, and the coach and coachee were afforded as much privacy as possible. Written consent for being recorded was obtained from all participants, and each was granted the right to withdraw at any time. Consequently, the participants were unaware of the actual aim of the research project,

which was a comprehensive study of multimodal communication in coaching (including the analysis of gestural metaphor). Furthermore, participants were afforded the opportunity to engage in discourse with the coach prior to and following each session, without being recorded, thus enabling them to address any concerns or receive additional feedback.

The coaches who recorded sessions for this study underwent training in accordance with the standards set forth by the International Coaching Federation (ICF) and were certified as Associate Certified Coaches by the ICF. This certification attests to their competence in supporting individuals in their personal and/or professional development. Additionally, they received training and certification as Clean Language coaches in the Polish version of Clean Language. The coaches who were recorded as part of the Career Clean Coaching corpus were hired and remunerated as part of a funded research project. Each coachee who was recorded as part of the Career Clean Coaching corpus participated in two sessions. The average duration of a session in the Career Clean Coaching corpus was 40 minutes. While the topic was limited to career-related matters, the coachees were permitted to select any aspect of their professional careers. This could include their current circumstances, future aspirations, or their general understanding of professional careers in the context of other activities in their lives. A portion of the recordings was transcribed using ELAN software (Sloetjes & Wittenburg 2008), and the utterances of the coachees were utilized in the subsequent analyses. The resulting Career Clean Coaching corpus, which was employed in this study, encompasses 104,914 words from 23 speakers (coachees), 18 of whom are female and 5 of whom are male.

4. Results

4.1. Study 1: The Five senses of “I think” in Career Clean Coaching

The verb “to think” (“myślę” in Polish) is highly polysemous in Polish, as it is in many other languages. A study utilizing dictionary-based annotation enriched with cognitive linguistics concepts of agentivity and conducted on the Polish written corpus revealed varying degrees of agentivity in five distinct senses of the verb “to think” (Kokorniak & Konat 2012). Increasing the sense of agentivity is a primary objective for both coaches and coachees, and the usage of “I think” may indicate whether the speaker perceives themselves as an agent or an experiencer. The objective is to identify the same set of five senses in the Career Clean Coaching corpus and to focus on the first-person singular form of the verb “I think”.

A manual annotation was conducted on all 300 instances of “I think” in the Career Clean Coaching corpus, following the methodology introduced in Kokorniak & Konat (2012). The taxonomy employed here adheres to a cognitive view of language, establishing a continuum between the subject as an agent and the subject as an experiencer. The subject as an agent is the most active, volitional, controlling and responsible, corresponding to Sense 5. In contrast, the subject as an experiencer is more passive, reflecting and re-membering (Sense 1). In order to demonstrate the distinction between the five senses, we propose an English equivalent for each example, which we believe is the closest to

the given facet of the meaning conveyed by the speaker. At the same time, it should be noted that in English, the word "think" can also be used in a word-by-word translation.

Firstly, the use of "I think" allows speakers to express mental processes in which they are not responsible (such as the process of remembering) and non-volitional, as exemplified in Example 1 (Sense 1). An indicator of this sense is the construction "I think about" – "myślę o". An English equivalent of this meaning can be conveyed by the word "remember".

Example 1:
*no taką że czuję ulgę w momencie nawet kiedy sobie **myślę** o tym że ja tej kłody nie mam na rękach*
(particle).such.that.feel.relief.in.moment.even.when.1SGthink.about.this.that.I.this.log.not. have.on.hands
[I feel relieved in the moments **I think** that I no longer have this log on my hands.]

Secondly, the process of thinking may be described as more volitional, whereby the speaker is establishing relationships between different concepts in the process of imagination. This process is volitional, yet the speaker lacks complete control over it and is not solely responsible for the truth value of it, as evidenced by Example 2 (Sense 2). In English, this can be expressed with the verb "to imagine".

Example 2:
*nie wiem czy się będę w niej sprawdzać po prostu sprawdzić czy jest faktycznie tak jak **myślę** tak zweryfikować.*
no.know.if.(reflexive).will.in.it.perform.just.simple.check.if.is.really.just. how.1SGthink.like.verify

[I don't know if I will perform well in it, just to check if this is really how **I think** it will be, just to verify.] or [I believe that the best way to proceed is to verify this information.]

Thirdly, when "I think" precedes the proposition, it is employed to diminish the assertion that follows, which aligns with Sense 3 in our taxonomy. This usage is frequently accompanied by the use of the word "że" ("that") and generally corresponds to the English use of "I think" as a discourse marker used to hedge propositions, as shown in Example 3.

Example 3:
*ja **myślę** że to taki cel do którego gdzieś tam dążę to co chciałabym osiągnąć* I.think.that.it.such.aim.to.which.somewhere.there. strive.this.what.would.acheive
[**I think** that this is the goal towards which I am striving, something I am striving to achieve.]

Fourthly, more volitional usage is related to the expression informing about speakers' plans. In Polish, "to think" may be used to express the plan or intention to do something, as we can see in Example 4. (Sense 4).

Example 4:
*a ostatnio **myślę** czy nie zostać krawcową także to się naprawdę może wszystko wydarzyć*
and.lately.1SGthink.whether.not.become.tailor.so.it.(reflexive).really.can. everything happen
[And lately **I've been thinking** whether to become a tailor, so really anything can happen.]

Fifth, speakers of CCC also employed the phrase "I think" to describe the process of deliberating, considering, thinking deeply, and judging reality, as exemplified in Example 5 (Sense 5). In this instance, the subject is the most agentive, as she controls the process and is able to make conclusions with confidence. She is able to induce or stop the process at will.

Example 5:
*ma minę obojętnie ja nie wiem po prostu mam i od razu coś sobie **myślę** na przykład*
has.expression.nevermind.I.do.not.know.just.(I)have.and.instantly.something. **(I)think**.
for.example *w związku z tym aha to on pewnie tak albo taka* in.relation.with.this.aha.
so.he.surely.(is)so.or.(she)is(so).
[Someone is making a face and I – I don't know, this is how I am – and instantly **I am thinking** on the base of it: she is probably like that and he is like that.]

The results of manual annotation of the five senses of the phrase "I think" in the Career Clean Coaching corpus revealed that the most prevalent sense was "Sense 3: to believe in the truth of a proposition". This sense constitutes 91% of all instances in the corpus (see Table 1).

Table 3: The distribution of the five senses of "I think" in the Career Clean Coaching corpus

Sense	Sense description	English equivalent	Number of instances
Sense 1	to keep in mind	remember	4
Sense 2	to find relationships among things	imagine	3
Sense 3	to believe in the truth of a proposition	think	**274**
Sense 4	to intend to do something	plan	6
Sense 5	to deliberate, to consider, to think deeply	judge	13
Total instances			300

The results indicate that the most prevalent usage of the phrase "I think" by speakers of CCC is the moderate level of agentivity (Sense 3). Coachees rarely refer to the most agentive and reflective sense 5, where the responsibility for the judgement is high. Conversely, they do not adopt the opposite end of the continuum, as the usage of Sense 1 is not frequent. The reason for this is that the topic under discussion is abstract and complex, namely future career and professional development. Speakers in CCC are not yet prepared to assume full responsibility for their actions, as expressed by the phrase

"I think", and therefore tend to utilize Sense 3 to mitigate the agentivity of their utterances.

The multiple senses of the Polish expression "I think" reflect the multifunctionality of this utterance. As observed by Danielewiczowa (2002), when "I think" in Polish precedes the proposition (often in the pattern "myślę, że p" – "I think that p"), it serves the function of a hedge, softening the commitment to the claim which follows. Danielewiczowa's definition aligns with our understanding of Sense 3. This indicates that in the CCC corpus, the expression "I think" is employed primarily in the capacity of a hedge, signifying the coachees' lack of certainty. This is consistent with the observation of discourse markers indicating hesitation and uncertainty in psychotherapeutic sessions. The presence of discourse markers indicating uncertainty suggests that coaching sessions are similar to psychotherapeutic sessions, as predicted by our application of Ferrara's model.

4.2. Study 2: The frequency of "I think" in Career Clean Coaching

The objective of this study is to ascertain the frequency of "I think" in the Career Clean Coaching corpus and to compare it with the frequency in other corpora. In order to provide a reference point for the analysis, three corpora of Polish language were utilised. The NKJP, ORIGAMI and PELCRA corpora were used, which represent spoken language.

The NKJP (National Corpus of Polish, Przepiórkowski et al. 2010) is a balanced corpus, comprising both written and spoken language, with a greater prevalence of the former. It contains over 300 million tokens. The ORIGAMI corpus (Karpiński et al. 2008) comprises 9,239 tokens from 40 speakers engaged in task-oriented dialogues (folding paper figures). The spoken data employed in this study was derived from the Polish and English Language Corpora for Research and Applications (PELCRA) (Pęzik 2012). This comprised collections of spoken language, dialogues and polylogues. The topics covered in this corpus are typically informal, including discussions on holiday planning, health, and politics. The participants in the PELCRA corpus provided their consent for the recording of their conversations prior to the commencement of the recording process. However, they were not informed of the precise moment at which the recording would commence.

Table 4: Comparison of frequencies of "I think" in selected corpora of Polish language

Corpus type	Corpus name	Total number of tokens in corpus	Frequency of "I think"	Frequency of "I think" per 10 000 tokens
balanced	**NKJP**	300 000 000	1000	0,03
spoken	**PELCRA**	2 126 961	1 119	5,26
	ORIGAMI	9 239	5	5,41
	Career Clean Coaching	104 914	300	28,59

Table 4 presents the frequencies of the verb “I think” in the compared corpora. Given the varying availability and size of spoken and written corpora, the size of our samples also varies. Consequently, the frequency was reported as the number of occurrences per 10,000 words. In the balanced version of the National Corpus of Polish, for the total sample size of 300 million words, we can observe only 1000 instances of the verb “I think”, which gives us a normalized value of 0.03 per 10,000. This indicates that the verb is exceedingly uncommon in this corpus, given that the balanced sample is still predominantly composed of written language. The frequent use of the verb “I think” is a distinctive feature of spontaneous spoken discourse.

In order to ascertain whether the lexical choices of speakers in career coaching sessions are indeed specific to this discourse, or merely a typical feature of spoken language, we conducted a comparative analysis of the CCC corpus with two other types of spoken corpora: The spoken corpora under consideration are PELCRA and ORIGAMI. As illustrated in Table 4, the relative frequency of the verb “I think” in both the compared spoken corpora is comparable, with a value of approximately 5. This suggests that the frequency of use of “I think” remains consistent across both standard, everyday conversations (including dia- and polylogues) and task-oriented dialogues. Nevertheless, the frequency of occurrence of the verb “I think” increases in the context of career coaching sessions, with a rate of 28.69 per 10,000 tokens.

To ascertain whether the observed differences between career coaching sessions and other conversations were statistically significant, we conducted a comparative analysis of the frequencies of the phrase “I think” in the three spoken corpora with the Log Likelihood (LL) value. Comparisons of the CCC corpus (coaching sessions) with the PELCRA (conversational spoken Polish) and ORIGAMI (task-oriented dialogues) revealed statistically significant differences (LL value 478.31, $p < 0.0001$ and LL value 24.75, $p < 0.0001$, respectively). However, the comparison between two other spoken corpora (PELCRA and ORIGAMI) revealed no statistically significant differences in terms of the use of the phrase “I think”.

4.3. Study 3: Speaker differences in Career Clean Coaching

As previous research suggested, variability of most features in spoken corpora is high, and aggregated numbers should be reconfirmed with tests taking between-speaker variability into account (Brezina & Meyerhoff 2014). The data also reflected this high degree of between-speaker variability, as illustrated in Figure 1. The histogram shows the variability in the frequency of the utterance “I think” (normalized per 10,000 tokens) between speakers. Five speakers uttered the phrase between one and five times, with the lowest number being 4.82. Seven speakers uttered the phrase between ten and twenty times, and four speakers uttered it between twenty and thirty times. As the data is highly right skewed, the remaining values are unequally distributed. The highest frequency, found in only one speaker, is 243 instances (normalized per 10,000).

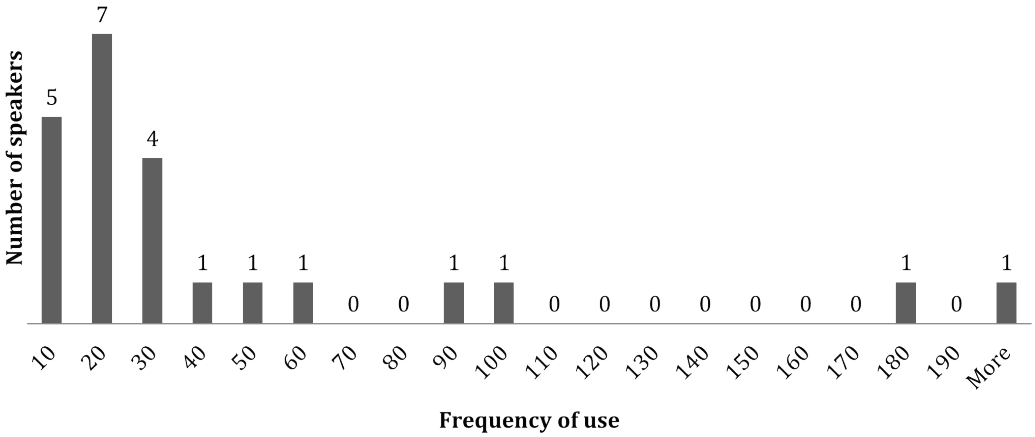


Figure 1: The distribution of frequencies (normalized per 10 000 tokens) of "I think" between speakers in Career Clean Coaching corpus.

To account for the high between-speaker variability, intergroup comparisons were performed using the Mann–Whitney U Test analysis (speakers from CCC versus PELCRA speakers). In order to obtain a reference corpus of a similar size and characteristics to the target (CCC) corpus, a systematic sample was created from PELCRA[1]. Currently, PELCRA does not allow for global identification of speakers, thus it is not possible to determine which utterances belong to a particular speaker. However, for each utterance, data about gender, age and education level are provided in the PELCRA results. Consequently, in order to create a corpus with a similar size and gender balance to that of the CCC corpus, it was necessary to create 23 artificial 'speakers'. The PELCRA sample was divided into 23 'speakers' based on age, gender and education. This resulted in 18 female speakers and 5 male speakers, all aged between 25 and 35 and with a higher education degree. This data was used to conduct a Mann Whitney U Test (see Table 5).

Table 5. A comparison of the frequency of the use of the phrase "I think" between speakers in the CCC and Pelcra corpora is presented in Table 5. The highest frequency of the use of "I think" among female speakers is as high as 243.68 instances per 10,000. However, this speaker was an outlier. As illustrated in Figure 1, the histogram reveals that only two speakers (both female) in the total CCC corpus achieved a value above 100 instances per 10,000 tokens. The second speaker attained a value of 173.61. In the context of everyday conversations (PELCRA corpus), the highest observed value was 19.96 instances per 10,000 (for a female speaker), with the next speaker (male) exhibiting

[1] It was not possible to preserve gender balance in our group due to the lack of male participants volunteering in coaching sessions. This is why for the Mann–Whitney U Test test we created the sample with equally high number of females. In this group the effect of the pragmatic setting (being in the coaching session or conversation) is still statistically significant.

Table 5: Comparison of the frequency of the use of "I think" between speakers in CCC and Pelcra corpus

		Career Clean Coaching			PELCRA		
Speaker ID	Speaker Gender	Words	Sum of "I think"	Frequency per 10 000	Words	Sum of "I think"	**Frequency per 10 000**
1	Female	5344	11	20.58	3107	3	9.66
2		1055	3	28.44	4157	5	12.03
3		1440	25	173.61	3999	1	2.50
4		3380	6	17.75	5135	2	3.89
5		5246	4	7.62	6538	1	1.53
6		3458	14	40.49	4682	4	8.54
7		11154	9	8.07	4199	2	4.76
8		11202	34	30.35	5668	2	3.53
9		1489	2	13.43	6139	2	3.26
10		6767	67	99.01	3700	4	10.81
11		2726	22	80.70	8285	13	15.69
12		996	1	10.04	3471	1	2.88
13		4016	7	17.43	4168	2	4.80
14		1108	27	**243.68**	3996	3	7.51
15		5871	30	51.10	4370	1	2.29
16		2289	5	21.84	4709	0	0.00
17		1948	3	15.40	4509	9	**19.96**
18		6417	12	18.70	4328	2	4.62
19	Male	1381	4	28.96	4185	2	4.78
20		3617	3	8.29	3560	7	19.66
21		1996	2	10.02	3559	2	5.62
22		12454	6	4.82	3899	3	7.69
23		3879	3	7.73	1498	0	0.00
SUM		99233	300	**(Av.)30.23**	101861	71	(Av.)6.97

a value that was very close to this at 19.66. Figure 2 presents the distribution and outliers between the two groups of speakers. The distribution of the frequency of the utterance "I think" between 23 speakers in the Career Clean Coaching Corpus (left) and the sample of spoken Polish (PELCRA corpus – right) is presented in Figure 2. This figure shows the number of instances of the utterance per 10,000 tokens.

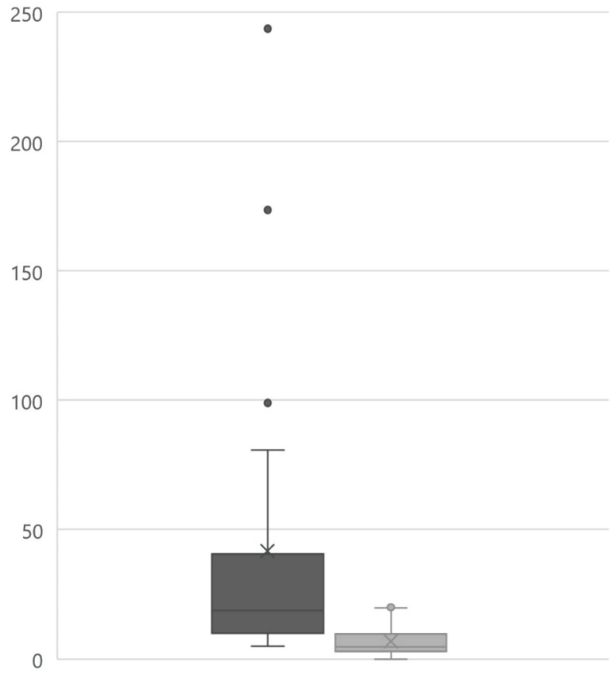


Figure 2: Distribution of the frequency of "I think" between 23 speakers in the Career Clean Coaching Corpus (left) and the sample of spoken Polish (PELCRA corpus – right), normalized per 10,000 tokens

To account for the high between-speaker variability observed in this data, intergroup comparisons were performed using the Mann–Whitney U Test analysis (speakers from CCC versus PELCRA speakers). The U-Mann Whitney test yielded statistically significant differences between two groups of speakers, returning U-values of 61 and 100, respectively, for normalized per 10,000 tokens and actual values. The Z-test was significant at $p < 0.01$, indicating that the two groups of speakers differed with regards to their frequency of use of "I think", despite between-speaker variability.

5. Discussion

The analysis presented in this paper provides evidence that speakers in specific dialogical situations – career coaching sessions – use the expression "I think" more frequently than similar speakers in task-oriented dialogues or in everyday conversations. Furthermore, the majority of instances of "I think" are used in the function of a hedge, softening the assertive force of the proposition. This allows us to claim that, in addition to individual factors already recognised in the literature, such as gender or age, contextual and pragmatic factors, namely the topic and setting of the conversation, may also influence the use of "I think". This finding is in accordance with previous studies (Mullan

2010, Ponterotto 2014, Helasvuo 2014, Kärkkäinen 2003, Brezina 2013) – as we found that the frequency of the use of "I think" is highly variable among speakers and may be considered a part of an individual speaker's linguistic repertoire. Although speakers vary in the frequency of their use of "I think", in the specific context of coaching sessions, they use this expression at a much higher rate than in other situations. We argue that the conversational setting, along with the specific protocol and abstract and difficult topic, influence speakers' lexical choices. Additionally, the observed frequency of the use of "I think" as a hedge suggests that this is a distinctive feature of career coaching sessions. During career coaching sessions, speakers tend to grapple with complex and abstract topics such as future career, beliefs and attitudes. We posit that, akin to the metaphor, the frequent use of "I think" in career coaching sessions serves as a linguistic tool for speakers to conceptualize complex subjects, and to signal their uncertainty. This paper analyses one discourse marker of coaching dialogues. Further studies are required to fully understand the nuances of natural, spontaneous dialogue, such as coaching sessions. Individuals discussing abstract concepts, such as their future or career, often feel uncertain and find it challenging to articulate their thoughts. In addition to hedges, other linguistic devices, such as metaphors, should be recognized and analysed as indicators of abstract and complex conversations during coaching sessions.

Conflict of Interest: The authors declare no conflict of interest
Authors' Contributions: The authors confirm sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Acknowledgements

The paper has been prepared within a research project funded by the *Polish National Science Centre* (decision DEC-2011/03/D/HS6/05993). Authors would like to thank all team members and subjects of the study.

References

Adler, R.B. & Towne, N. 1996. *Looking out/looking in: Interpersonal communication*. Miasto: Wydawnictwo.
Baker, P. 2010. *Sociolinguistics and corpus linguistics*. Edinburgh: Edinburgh University Press.
Baumgarten, N. & House, J. 2010. *I think* and *I don't know* in English as lingua franca and native English discourse. *Journal of Pragmatics* 42(5). 1184-1200.
Biber, D. 2012. Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory* 8(1). 9-37.
Brezina, V. 2013. *Certainty and uncertainty in spoken language: In search of epistemic sociolect and idiolect*. Frankfurt am Main: Peter Lang.
Brezina, V. & Meyerhoff, M. 2014. Significant or random? A critical review of sociolinguistic generalisations based on large corpora. *International Journal of Corpus Linguistics* 19. 1-28. https://doi:10.1075/ijcl.19.1.01bre.
Cameron, L. & Maslen, R. 2010. *Metaphor analysis: Research practice in applied linguistics, social sciences and the humanities*. London: Equinox Pub.
Cutting, J. 2000. *Analysing the language of discourse communities*. Amsterdam etc.: Elsevier.
Danielewiczowa, M. 2002. *Wiedza i niewiedza: Studium polskich czasowników epistemicznych*. Warszawa: Katedra Lingwistyki Formalnej UW.

Egbert, M. & Yufu, M. & Hirataka, F. 2016. An investigation of how 100 articles in the *Journal of Pragmatics* treat transcripts of English and non-English languages. *Journal of Pragmatics* 94. 98-111.

Fabiszak, M. 2008. Corpus frequency as a guide to metaphor labelling. In Wąsik, Z. & Komendziński, T. (eds.), *Metaphor and cognition* (Special volume of *Philologica Wratislaviensia: From grammar to discourse*), 149-162. Frankfurt am Main: Peter Lang.

Ferrara, K.W. 1994. *Therapeutic ways with words*. New York–Oxford: Oxford University Press.

Freitag, R.M.K. & Cardoso, P.B. & Tejada, J. 2022. Linguistic and paralinguistic constraints on the function of *(eu) acho que* as DM in Brazilian Portuguese: A multilevel approach. *Pragmatics & Cognition* 29(2). 324-46. https://doi.org/10.1075/pc.21024.fre.

Górski, R. & Lewandowska-Tomaszczyk, B. (eds.). Narodowy Korpus Języka Polskiego. Warszawa: Wydawnictwo Naukowe PWN.

Helasvuo, M.L. 2014. Agreement or crystallization: Patterns of 1st and 2nd person subjects and verbs of cognition in Finnish conversational interaction. *Journal of Pragmatics* 63. 63-78.

Juszczyk, K. 2017. M*ądrość metafory: Komunikacyjny model metafory kognitywnej*. Poznań: Coachspace.

Kaltenböck, G. 2009. Initial *I think*: Main or comment clause. *Discourse and Interaction* 2(1). 49-70.

Kamasa, V. 2013. Naming "in vitro fertilization": Critical discourse analysis of the Polish Catholic Church's official documents. *Procedia – Social and Behavioral Sciences* 95. 154-159.

Kärkkäinen, E. 2003. *Epistemic stance in English conversation: A description of its interactional functions, with a focus on* I think. Amsterdam–Phialelphia: John Benjamins.

Karpiński, M. & Juszczyk, K. & Jarmołowicz-Nowikow, E. & Malisz, Z. & Szczyszek, M. 2008. Rejestracja, transkrypcja i tagowanie mowy oraz gestów w narracji dzieci i dorosłych. *Investigationes Linguisticae* 16. 83-98.

Kokorniak, I. & Konat, B. 2012. Agent or experiencer? A search for the subject role in the mental verb *myśleć* 'think' in Polish. In Kosecki, K. & Badio, J. (eds.), *Cognitive processes in language*, 77-86. Frankfurt am Main: Peter Lang.

Konat, B. & Juszczyk, K. 2015. Multimodal communication in career coaching sessions: Lexical and gestural corpus study. *Lodz Studies in Language* 37. 195-212.

Lakoff, G. 1975. Hedges: A study in meaning criteria and the logic of fuzzy concepts. In Hockney, D. & Harper, W. & Freed, B. (eds.), *Contemporary research in philosophical logic and linguistic semantics* (The University of Western Ontario Series in Philosophy of Science), 221-271. Dordrecht: Springer.

Lakoff, G. & Johnson, M. 2003. *Metaphors we live by*. 2nd edn. Chicago: University of Chicago Press.

Lawley, J.D. & Tompkins, P.L. & 2000. *Metaphors in mind: Transformation through symbolic modelling*. 1st edn. London: Developing Company Press.

Lewandowska-Tomaszczyk, B. 2011. Nowe wyzwania w jakościowej i ilościowej metodologii analizy języka. *Biuletyn Polskiego Towarzystwa Językoznawczego* 67. 51–71.

McCarthy, M. 1998. *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.

McMullen, L.M. & Conway, J.B. 1996. Conceptualizing the figurative expressions of psychotherapy clients. In Mio, J.S. & Katz, A.N. (eds.), *Metaphor: Implications and applications*, 59-72. Mahwah, N.J.: Psychology Press.

Mullan, K., 2010. *Expressing opinions in French and Australian English discourse: A semantic and interactional analysis*. Amsterdam: John Benjamins.

Oleskeviciene, G.V. & Liebeskind, Ch. 2021. Multiword expressions as discourse markers in Hebrew and Lithuanian. In Bizzoni, Y. & Teich, E. & España-Bonet, C. & van Genabith, J. (eds.), *Proceedings for the First Workshop on Modelling Translation: Translatology in the Digital Age*, 46-56. Association for Computational Linguistics. (https://aclanthology.org/2021.motra-1.5)

Pawelczyk, J. 2011. *Talk as therapy: Psychotherapy in a linguistic perspective*. Boston–Berlin: De Gruyter Mouton.

Pęzik, P. 2012. Wyszukiwarka PELCRA dla danych NKJP. In Przepiórkowski, A. & Bańko, M. & Górski, R. & Lewandowska-Tomaszczyk, B. (eds.), Narodowy Korpus Języka Polskiego, 253-279. Warszawa: Wydawnictwo Naukowe PWN.

Pieśkiewicz, B. & Kołodkiewicz, M. 2011. Metafora w coachingu: Zastosowanie metody Clean Coaching w pracy z organizacjami. In Czarkowska, L. (ed.), *Coaching: Katalizator rozwoju organizacji*, 207-223. Warszawa: New Dawn.

Ponterotto, D. 2014. The risks of uncertainty: Hedging strategies in rape trial discourse. Language and Dialogue 4. 93-111. https://doi:10.1075/ld.4.1.06pon.

Przepiórkowski, A. & Bańko, M. & Górski, R. & Lewandowska-Tomaszczyk, B. (eds.), *Narodowy Korpus Języka Polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Sato, S. 2017. On establishing I think as a final particle in interactions: Some comparisons with sentence-final particles in Japanese. *Journal of Pragmatics* 110. 83-98.

Sloetjes, H. & Wittenburg, P. 2008. Annotation by category-ELAN and ISO DCR. 6th International Conference on Language Resources and Evaluation (LREC 2008).

Sullivan, W. & Rees, J. 2008. *Clean Language: Revealing metaphors and opening minds*. 1st edn. Carmarthen, Wales: Crown House Publishing.

Tay, D. 2013. *Metaphor in psychotherapy: A descriptive and prescriptive analysis*. Amsterdam–Philadelphia: John Benjamins Publishing.

Thorpe, S. & Clifford, J. 2003. *The coaching handbook: An action kit for trainers and managers*. 1st edn. London: Kogan Page.

Tolvanen, E. 2024. 'I think' in Swedish L1 and L2 group interactions. *Pragmatics* 34(4). 615-641. https://doi.org/10.1075/prag.22060.tol.

Wang, Q. 2020. A corpus-based contrastive analysis of I think in spoken Hong Kong English: Research from the International Corpus of English (ICE). *Australian Journal of Linguistics* 40(3). 319-45. https://doi.org/10.1080/07268602.2020.1823817.

DOI: 10.14746/linpo.2024.66.1.2

Morphological causatives in Moroccan Arabic: An Optimality-Theoretic account

Ayoub Loutfi

Hassan II University of Casablanca
a.loutfi@flbenmsik.ma | ORCID: 0000-0001-7830-3173

Abstract

This study investigates the formation of the morphological causative in Moroccan Arabic within the framework of Optimality Theory. The pivotal process involved in this construction is morphological gemination, where the second consonant of the root is doubled (e.g., *ktəb* ~ *kəttəb*). A perplexing aspect of this process is its consistent targeting of the second consonant, while excluding ill-formed words such as **kktəb* and **ktəbb*. In this study, we argue that the linearization of the causative morpheme in MA is governed by phonological well-formedness interacting with the morphological process of causativization, resulting in infixation. This phenomenon exemplifies the Emergence of the Unmarked (McCarthy & Prince 1994; Alderete et al. 1999), where the otherwise inactive markedness constraint *COMPLEX$_{\text{ONSET}}$ in the language assumes explanatory prominence. The strength of the analysis proposed here lies in treating the infixal process as an outcome of simple and universal constraints, primarily motivated by demands on prosodic well-formedness, without recourse to language-specific templatic patterns. This approach highlights the implications of phonological constraints in capturing morphological processes and suggests that templatic patterns may emerge from phonological well-formedness rather than being innate to the language.

Keywords: morphological causatives, Moroccan Arabic, Optimality Theory, positional faithfulness, prosodic morphology

1. Introduction

In Moroccan Arabic (MA henceforth), causatives are formed through the infixation of a featureless consonantal mora to the verbal root (McCarthy 1993; Bennis 1992; Boudlal 2001)[1]. However, the consistent targeting of the second consonant in the gemination

[1] The data is taken from the urban variety spoken in Rabat-Sale (Coastal East, Central), a variety of which I am a native speaker. See section 2 for more details.

process, while excluding ill-formed words like *kktəb and *ktəbb, poses a puzzling question. The examples provided in (1) illustrate this phenomenon[2].

(1)

Perfective[3]		**Causative Form**	
ktəb	'he wrote'	*kə**tt**əb*	'to make someone write'
ħRəb	'he ran away'	*ħə**RR**əb*	'to make someone run away'
ʃRəb	'he drank'	*ʃə**RR**eb*	'to make someone drink'
xrəʒ	'he went out'	*xə**rr**əʒ*	'to make someone go out'
lʕəb	'he played'	*lə**ʕʕ**əb*	'to make someone play'

As illustrated in (1), the phonological elements involved in morphological causatives in MA are exclusively derived from the root, reflecting a base-dependence effect (Kager 1999). Additionally, the structure of the causative morpheme remains consistently invariant across the paradigm, comprising solely a consonantal mora whose phonetic properties vary as the second radical consonant of the root varies. Furthermore, the positioning of the causative morpheme relative to the root consistently skips over the root's first consonant.

This study aims to explain why gemination occurs infixally within the root. While various scholars have described the process of morphological causatives (see Harrell 1962 for instance), the rationale behind why this morpheme targets the second radical consonant of the root has been explained either through unsatisfactory stipulations (McCarthy 1993; Bennis 1992; Boudlal 2001) or by lacking empirical support. Two main viewpoints have emerged in this regard. The first posits the template as a morphological unit (McCarthy 1979, 1981, 1993; Bennis 1992). The second contends that the morphological process is triggered by the privileged status of the root's first radical consonant, which not only avoids phonological processes but also morphological ones (Noamane 2020).

The inadequacies of both the templatic and positional faithfulness approaches become apparent upon closer examination of their empirical grounding. First, while the templatic analysis may align with certain verb patterns, it fails to accurately capture the semantic and syntactic nuances of causatives. Moreover, it neglects the pivotal role of phonological well-formedness in triggering the infixal process, which is a critical factor in understanding the phenomenon (McCarthy 1993; Yu 2007).

Similarly, the positional faithfulness approach encounters challenges when confronted with empirical evidence highlighting other phonological processes that affect the ostensibly privileged position. These processes, such as long-distance consonant harmony and

[2] For the transcription of data in this paper, the International Phonetic Alphabet is adopted. Emphatic consonants are represented using capital letters, and geminate consonants are denoted by doubling the respective consonant symbol.

[3] MA, like many Arabic dialects, exhibits a distinct phonemic and vocalic inventory. Notably, in MA the presence of schwa [ə] is restricted and typically does not appear in open syllables. For further information on the phonological features and syllable structure of MA, interested readers may refer to works such as Benhallam (1990), Boudlal (2001), and Al Ghadi (2022).

secondary labial assimilation, disrupt the notion of a fixed positional hierarchy (Zellou 2010). Consequently, both approaches fail to provide a comprehensive explanation of morphological causatives in MA.

Hence, we advocate an alternative framework rooted in parallel Optimality Theory (OT), which offers a more robust and empirically grounded account. Within this framework, we demonstrate that the causative morpheme's appearance is constrained by *COMPLEX$_{ONSET}$, explaining its infixation relative to the root. Crucially, the dominance of *COMPLEX$_{ONSET}$ underscores MA's allowance for complex onsets, aligning with OT's concept of the Emergence of the Unmarked (TETU) (McCarthy & Prince 1994; Alderete et al. 1999).

By adopting the principles of OT and emphasizing the interplay between phonological and morphological constraints, our analysis provides a more nuanced understanding of morphological causatives in MA, transcending the limitations of previous approaches.

The remainder of this paper is constructed as follows: Section 2 focuses on the data source. In section 3, we examine the MA syllable structure to understand its influence on morphological processes. Section 4 introduces OT and the concept of TETU, establishing the theoretical foundation. Section 5 reviews previous analyses of MA causatives, including templatic analysis and positional faithfulness approaches, and evaluates their strengths and weaknesses. Section 6 presents the proposed analysis within the OT framework, outlining the constraints and mechanisms governing causative formation and discussing the theoretical implications. Finally, section 7 concludes the study by summarizing the key findings.

2. Data source

This study primarily uses data from the urban MA variety spoken in Rabat-Sale, a region in the Coastal East Central region of Morocco. Broadly, MA is classified as a member of the Hamito-Semitic language family. It is often referred to as Afro-Asiatic, a broader category encompassing languages such as Berber (Amazigh), Semitic, Cushitic, Egyptian, and Chadic. Languages within this family share several linguistic features, notably non-concatenative morphology, where word formation primarily occurs through internal modifications of the root rather than simple morpheme concatenation, as observed in languages like English or Italian (McCarthy 1976, 1981). Other distinctive features include broken plurals in nouns (Al Ghadi 2014), emphatic and glottalized consonants, and intricate verb inflection and derivation processes.

Despite their linguistic relatedness, Standard Arabic and MA exhibit significant differences across lexical, phonological, morphological, and semantic domains (Benhallam and Dahbi 1990; Boudlal 2001). Additionally, MA serves as the primary spoken language in Morocco, lacking codification or standardization, and is exclusively used in oral communication. Because of language contact, MA has absorbed numerous lexical items from Amazigh and foreign languages, exemplifying its dynamic linguistic evolution (Boudlal 2001).

MA encompasses various regional varieties, each of which exhibits intradialectal variations. These include the urban, mountain, bedouin, and Hassani varieties. However,

despite these variations, all MA varieties are believed to share the same morphological process for deriving morphological causatives[4].

3. Moroccan Arabic syllable structure

In the investigation of the MA syllable structure, crucial insights emerge regarding its phonological organization, shedding light on the language's intricate phonotactic patterns. As documented in previous studies (Benhallam 1990; Boudlal 2001; Bensoukas & Boudlal 2012; Al Ghadi 2022), MA exhibits a vowel inventory consisting of three basic vowels [i, u, o] along with an epenthetic schwa [ə]. The epenthetic nature of schwa is attributed to its restricted distribution, which is notably absent in open syllables. This phenomenon is evidenced in the data presented, where schwa appears in positions dictated by phonotactic constraints, such as between the first two root consonants when the third-person plural affix is added.

(2)

ktəb	'he wrote'	*kətb-u*	'they wrote'
DRəb	'he hit'	*DəRb-u*	'they hit'
gləs	'he sat'	*gəls-u*	'they sat'
ʒbəd	'he pulled'	*ʒəbd-u*	'they pulled'
ħRət	'he plowed'	*ħəRt-u*	'they plowed'
qtəl	'he killed'	*qətl-u*	'they killed'

Furthermore, the MA syllable structure encompasses two primary types: CV and CVC, with additional forms derived from these basic structures (Benhallam 1990).

Of particular significance to the main argument of this paper is the observation that MA permits word-initial consonant clusters, as demonstrated across various linguistic categories such as verbs, adjectives, and nouns. Consider the examples below:

(3)

Verbs		**Adjectives**		**Nouns**	
ktəb	'he wrote'	*ʕrəʒ*	'lame'	*rʒəl*	'leg'
ʃTəħ	'he danced'	*ħwəl*	'cross-eyed'	*qfəz*	'cage'
DRəb	'he hit'	*kħəl*	'black'	*ʕsəl*	'honey'
gləs	'he sat'	*byəD*	'white'	*nməl*	'ants'

The allowance of complex onsets in MA aligns with the dominance of markedness constraints over constraints prohibiting complex onsets, as elucidated within the OT

[4] For an analysis of morphologically-derived causatives in Classical Arabic, see Loutfi (2024b). This study examines how morphological strategies like gemination, prefixation, and ablaut are governed by Contextual Allomorphy.

framework. In OT, allowing complex onsets amounts to saying that the constraint militating against complex onsets, namely *COMLEX$_{ONSET}$, is crucially dominated. This is apparent when the constraint interacts with the phonological constraint active in the language[5]. Notably, while the *COMPLEX$_{ONSET}$ constraint may seem inactive due to the language's tolerance of complex onsets, its influence becomes apparent in cases where faithfulness constraints fail to determine optimal surface forms. Unless otherwise indicated, this study adopts Boudlal's (2001) OT treatment of syllable structure in MA:

(4)

a. *Constraints:*

***Min-σ:** Minor syllables are prohibited.

MAX$_{IO}$: Every segment present in the input form of a linguistic expression must have a corresponding segment in the output form. This constraint prevents deletion, ensuring that all elements of the input are retained in the output.

DEP$_{IO}$: Every segment present in the output form must have a corresponding segment in the input form. This constraint prevents the addition of epenthetic material, ensuring that no new segments are introduced in the output that were not present in the input.

***COMPLEX$_{ONSET}$:** More than one consonant in the onset position is prohibited.

b. *Ranking*:

Input : /bka/			MAX$_{IO}$	*Min-σ	DEP$_{IO}$	*COMPLEX$_{ONS}$
	a.	bə.ka			*	
☞	b.	bka				*
	c.	bk	*!			
	d.	b.ka		*!		

The *Min-σ constraint stipulates that minor syllables, defined as those containing fewer than two moras, are prohibited in the output forms of linguistic expressions. This constraint reflects a phonotactic preference against syllables that lack sufficient weight, ensuring that they are adequately structured and phonologically well-formed. Furthermore, MAX$_{IO}$ dictates that every segment present in the input form of a linguistic expression must have a corresponding segment in the output form. This constraint promotes faithfulness to the input, ensuring that all phonetic material present in the input is preserved in the output without deletion or alteration. DEP$_{IO}$, on the other hand, requires that every segment in the output must have a corresponding segment in the input, thereby militating against the addition of epenthetic material. Unlike MAX$_{IO}$, which is concerned with

[5] This paper refrains from delving deeply into the intricate details of the syllable structure in MA because such discussions would divert attention from the primary focus. Therefore, only the phonological aspects relevant to the current investigation are addressed. For instance, the undominated constraint prohibiting onsetless syllables was omitted, as its exclusion did not significantly impact the analysis. Interested readers seeking further exploration of these topics are encouraged to consult works such as Benhallam (1990), Boudlal (2001), and Al Ghadi (2022) and the references provided therein.

preventing deletion, DEP_{IO} enforces faithfulness to the input by prohibiting epenthesis, ensuring that no new segments are introduced in the output that were not present in the input.

The $*COMPLEX_{CODA}$ constraint prohibits the occurrence of more than one consonant in the coda position of a syllable in the output forms of linguistic expressions. This constraint reflects a phonotactic preference for simple syllable structures, ensuring that codas remain phonologically well-formed and avoiding complexity in syllable-final consonant clusters.

As evident from the tableau in (4-b) above, candidates (a) and (c-d) are dismissed due to violations of the higher-ranked constraints DEP_{IO} and MAX_{IO}, respectively. Conversely, candidate (b) remains faithful to the input and avoids contravening any undominated constraints, thereby emerging as the optimal surface form. Notably, $*COMPLEX_{ONSET}$ does not influence the selection or rejection of candidates in this instance because the language permits complex onsets, a situation that arises from $*COMPLEX_{ONSET}$ being low-ranked and dominated by higher-ranked faithfulness constraints. However, this constraint becomes crucial in contexts in which faithfulness constraints are less decisive, demonstrating its role in shaping surface forms. This interaction exemplifies Optimality Theory's concept of TETU (The Emergence of the Unmarked), where a typically inactive constraint can influence the output under specific conditions. This phenomenon is further explored in the following section.

4. OT and TETU

OT (Prince & Smolensky 2004) provides a constraint-based framework for analyzing phonological phenomena. A key concept in OT is TETU, which captures the observation that reduplicants tend to favor unmarked structures over marked structures in a language (McCarthy & Prince 1994). This section explores the theoretical underpinnings of TETU and its explanatory power in the OT context.

In OT, markedness refers to the relative complexity or naturalness of a linguistic structure. Unmarked structures are considered simpler and more universal, adhering to the principles of ease of articulation and learnability. However, marked structures deviate from these tendencies. For instance, single consonants are unmarked compared to consonant clusters, which require more complex articulatory gestures.

OT considers a set of universal constraints to evaluate the candidate outputs generated by GEN. Markedness constraints prioritize unmarked structures, whereas faithfulness constraints favor the preservation of underlying structures. These two types of constraints are often in conflict; in some instances, faithfulness constraints dominate markedness constraints, whereas in other cases, the opposite occurs. The concept of arises from the dynamic interaction between these constraints.

Consider a language in which syllables can generally have consonant codas, which are marked structures. Now, imagine a scenario where a high-ranked markedness constraint against complex codas dominates a lower-ranked faithfulness constraint that enforces identity between the root and the reduplicant in reduplication. In this case, the

high-ranked markedness constraint would prevent the emergence of codas in the reduplicated form, resulting in simpler, unmarked syllables without codas. TETU explains how the high ranking of the markedness constraint leads to avoidance of the marked structure (coda) in reduplicated words, despite the presence of a lower-ranked faithfulness constraint.

In fact, TETU holds significant theoretical weight within OT. First, it highlights the power of even low-ranked constraints. Although dominated by faithfulness constraints, markedness constraints can still influence outputs when faithfulness is inapplicable. This suggests that all constraints might be universally present in grammar even if they are not always active.

Overall, the proposed TETU provides a framework to explain how reduplication often simplifies underlying structures. It is the high-ranking markedness constraints, rather than faithfulness constraints, that push the reduplicant towards unmarked structures. This dynamic offers a powerful tool within OT for understanding why reduplicants tend to avoid marked structures. TETU sheds light on the interplay of universal constraints and their interaction with specific morphological processes like reduplication, contributing to a richer understanding of how languages function (Urbanczyk 2007).

5. Previous accounts of morphological causatives in MA

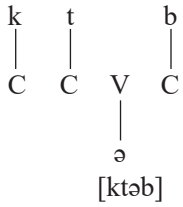
5.1. Templatic approach

Earlier approaches to morphologically derived causatives, influenced by McCarthy (1979, 1981), have posited a framework rooted in Autosegmental Phonology, as initially proposed by Goldsmith (1976). McCarthy's model suggests that Arabic verbs are structured across three distinct tiers within the lexical representation. First, the root tier contains the core verbal lexeme, which is composed of consonantal roots. Second, the skeletal tier, known as the prosodic template, establishes the canonical shape associated with specific meanings and grammatical functions, such as causativity, which is often represented by templates like CVCCVC. Finally, the vocalic melody tier encodes grammatical nuances like voice, aspect, and mood.

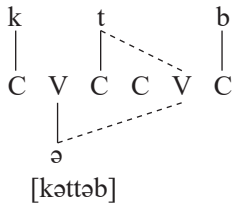
Within this framework, the derivation of causatives hinges on identifying the appropriate prosodic template and adhering to association conventions to fill vacant slots through consonant and vowel spreading mechanisms. Consequently, the causative forms emerge through a systematic mapping process that ensures a coherent alignment of consonant and vowel elements within the established prosodic template. Under this approach, the causatives are derived as follows:

(5)

a. Perfective form



b. Causative form



In both instances presented in (5), the consonantal root "ktb" and the vocalic melody are systematically associated with their respective slots. This association is governed by the Well-Formedness Constraint, which dictates that association lines must not cross and that, by the end of the derivation, all melody elements must be appropriately associated, with any unassociated materials being erased. The association mechanism follows a left-to-right directionality, ensuring a one-to-one mapping between the three tiers—root, prosodic template, and vocalic melody—without overlap.

As illustrated in (5), the distinction between the forms in (5-a) and (5-b) arises from the addition of a medial consonant slot. This addition triggers medial gemination, which is achieved through the automatic spreading of the medial consonant /t/ to fill the empty slot C. Consequently, the phonological content of the causative morpheme affix is realized by copying the phoneme melody of the root. Proponents of this view, such as Bennis (1992) and Bennis & Lazzi (1995), argue that morphological processes like infixation are driven primarily by the need to satisfy the prosodic template.

However, this approach faces two empirical challenges. First, the assumption that each template corresponds to a distinct meaning is not universally supported, as there exist verbs with template patterns that resemble those of causatives but lack causative properties. Consider the examples in (6):

(6)

*fə**ll**əħ*	'he farmed'
*sə**ll**əf*	'he lent (money)'
*wə**ll**əf*	'he accustomed'
*Tə**ll**əq*	'he divorced'
*Sə**RR**əf*	'he gave change'

Similar to morphological causatives, the verb class described in (6) exhibits the phenomenon of medial gemination and shares the same templatic shape as the causatives outlined in (5). This similarity becomes apparent when we consider the nominal counterparts of these verbs. As demonstrated in (7), their nominal forms are realized with a singleton consonant, mirroring the pattern observed in the causatives:

(7)

Nouns		**Verbs**	
flaħa	'farming'	*fəlləħ*	'he farmed'
səlf	'loan'	*səlləf*	'he lent (money)'
wəlf	'getting used to'	*wəlləf*	'he accustomed'
Tlaq	'divorce'	*Təlləq*	'he divorced'
SəRf	'cash change'	*Sərrəf*	'he changed money'

Moreover, as discussed in Benmamoun (1991) and Loutfi (2014, 2017, 2020, 2024a), one property of the causative morpheme in MA is that it functions as a valency-increasing morpheme that adds an agent argument to the clause. By way of illustration, consider the following examples:

(8)

a. *ʒamal* *xrəʒ*
Jamal go.out.PST[3M.SG]
'Jamal went out.'

b. *ʒamal* *xərrəʒ* *d-drari*
Jamal go.out\CAUS.PST[3M.SG] DEF-children
'Jamal made the children go out.'

c. *d-dərri* *lʕəb* *l-kuRa*
DEF-boy play.PST[3M.SG] DEF-ball
'The boy played football'

d. *ʒamal* *ləʕʕəb* *d-dərri* *l-kuRa*
Jamal play\CAUS.PST [3M.SG] DEF-boy DEF-ball
'Jamal made the boy play football.'

However, this is not the case for the class of verbs in (7) because these verbs do not alternate, as evidenced by the ambiguous and ungrammatical sentences in (9):

(9)

a. **Hiʃam* *səlləf* *(*ʒamal)*
Hicham lend.PST[3M.SG] (*Jamal)
'Hicham lent Jamal money'

b. *Hiʃam SəRRəf (*l-flus)
Hicham give.change.PST[3M.SG] (*the-money)
'Hicham gave change'

The second issue with the templatic approach pertains to the element associated with the V-slots. According to the original argument, the V-slots encode grammatical categories such as tense, aspect, and mood (see Bahloul 2008 for example). However, MA has undergone the loss of stem vowels. This becomes apparent when we compare the same class of verbs in both Standard Arabic (SA) and MA:

(10)

SA	**MA**	
katab	*ktəb*	'he wrote'
xaraʒ	*xrəʒ*	'he went out'
raħal	*rħəl*	'he moved out'
ħarab	*ħrəb*	'he ran away'
ʃarib	*ʃRəb*	'he drank'

In all MA verbs, the vowels encoding the active voice, namely /a…a/, are lost (Benhallam 1990; Al Ghadi 2022; Boudlal 2001). Associating schwas with V slots, as in (5), would treat them on a par with full vowels. This move is problematic because the restricted nature of the insertion of schwa in MA renders its status as purely epenthetic, breaking up clusters of consonants that the language does not tolerate. Therefore, the existence of forms whose derivations resemble those of morphologically derived causatives and the loss of stem vowels suggests that the prosodic template is not sufficient to identify morphologically derived causatives.

Theoretically, the templatic account overlooks the interaction between phonology and morphology in deriving the process of morphological gemination (Ussishkin 2000, 2007). For instance, the absence of ill-formed words such as **llaʕb* and **lʕabb* in the system is stipulated, rather arbitrarily, by the left-to-right association. However, we demonstrate that the process is derived through a joint consideration of both the phonological constraints active in the language, such as its syllable structure and the morphological process per se.

Constructing the process of morphological causatives as an instance of TETU, we illustrate that this process fundamentally arises from universal constraints, highlighting how phonological well-formedness influences morphological processes. Before we elaborate on our analysis, we review and argue against another account couched within the theory of OT.

5.2. Positional faithfulness

The positional faithfulness theory is based on three fundamental assumptions. First, privileged positions permit a wide range of marked segments, unlike unprivileged posi-

tions. This phenomenon is evident in languages with a rich consonantal system, where only a limited, ostensibly unmarked subset of segments/features are allowed in the featural content of affixes. Second, privileged positions trigger phonological processes, and third, they resist otherwise regular processes in the language (Yip 1991; McCarthy & Prince 1999; Beckman 1998, 2004; Alderete 2001; Lombardi 2004, among others). In MA, numerous contexts exhibit positional faithfulness. For example, the definite article affix /l/ regressively assimilates into the immediately adjacent coronal sound of the first radical consonant of the root (11-a). Otherwise, the /l/ morpheme is realized. The data below illustrate this general process:

(11)

UR			
a.	*/l-DaR/*	*DDaR*	'the house'
	/l-Suq/	*SSuq*	'the market'
	/l-tuma/	*ttuma*	'the garlic'
	/l-ʒlbana/	*ʒʒəlbana*	'the green peas'
	/l-ʃms/	*ʃʃəms*	'the sun'
	/l-ng^{w}ir/	*nng^{w}ir*	'the nagging'
	/l-Rajb/	*RRajəb*	'the churned milk'
b.	*/l-bab/*	*lbab*	'the door'
	/l-flfla/	*lfəlfla*	'the peppers'
	/l-qRfa/	*lqəRfa*	'the cinnamon'
	/l-klma/	*lkəlma*	'the word'
	/l-ʕRaDa/	*ləʕRaDa*	'the invitation'
	/l-ɦmm/	*lɦəmm*	'the worry'

As indicated by the data above, elements in the root, particularly the initial consonants, exhibit behaviors that trigger and resist the assimilatory process. These privileged behaviors are derived from the following ranking schema for positional phonological asymmetries:

(12)

IDENT-Position (F) >> C >> IDENT (F)
(Beckman 1998: 9)

Here, "C" represents the intervening markedness constraints. By dominating the non-privileged position faithfulness constraint IDENT (F), C implies neutralizing of contrast in this context. According to OT, assimilation is generally viewed as an instance of the markedness constraint AGREE (Lombardi 2004). In this case, the intervening markedness constraint is AGREE-Coronal, which competes with the positional constraints IDENT-ROOT (F) and IDENT-IO (F).

(13)

a. *Constraints:*

AGREE-CORONAL: Coronal clusters should agree in place.
IDENT-IO (F): Underlying featural specifications should remain the same.
IDENT-ROOT (F): The underlying feature specifications of the root must be the identical.

b. *Ranking:*
IDENT-ROOT (F) >> AGREE-CORONAL >> IDENT-IO (F)

At the core of this framework is the IDENT-ROOT (F) constraint, which prioritizes the preservation of the underlying feature specifications of the root. This constraint underscores the importance of maintaining the integrity of the root structure throughout the derivation process. Placing IDENT-ROOT (F) at the top of the ranking schema highlights its pivotal role in ensuring that the phonological properties of the root are faithfully preserved in the derived form.

In contrast, the AGREE-CORONAL constraint addresses the need for coronal clusters to agree in place. While crucial for maintaining phonological harmony within the word structure, AGREE-CORONAL is positioned below IDENT-ROOT (F) in the ranking schema. This placement suggests that while coronal agreement is important, it is subordinate to preserving the featural specifications of the root. Consequently, the analysis prioritizes the integrity of the root's phonological structure over coronal cluster agreement.

Finally, the IDENT-IO (F) constraint focuses on maintaining consistency in the underlying featural specifications across different morphemes. Positioned last in the ranking schema, IDENT-IO (F) underscores the importance of preserving featural content but indicates that this consideration is secondary to ensuring agreement in coronal features as dictated by AGREE-CORONAL. By placing IDENT-IO (F) below AGREE-CORONAL, the analysis suggests that while consistency in featural specifications is relevant, it is less crucial than achieving coronal feature agreement across segments.

(14)

IDENT-ROOT (F), AGREE-CORONAL >> IDENT-IO (F)

Input: /l-DaR/	IDENT-ROOT (F)	AGREE-CORONAL	IDENT-IO (F)
a. l-DaR		*!	
b. l-laR	*!		
☞ c. D-DaR			*

The application of positional faithfulness constraints in the analysis of phonological processes yields insights into the systematic nature of phonological asymmetries. In the presented example, the constraint AGREE-CORONAL plays a pivotal role in determining the optimal candidate by enforcing agreement within coronal clusters. Candidate (a) is deemed unviable due to its violation of AGREE-CORONAL, illustrating the constraint's influence on phonological decision-making. Meanwhile, candidate (b) satisfies AGREE-CORONAL but is ultimately ruled out by a higher-ranked positional constraint, highlighting the hierarchical nature of the constraint interaction. Candidate (c) emerges as the optimal choice, aligning with the dominating constraints and exemplifying the effectiveness of positional faithfulness in phonological analysis.

Moreover, the phenomenon of glide formation in MA provides additional evidence for the relevance of positional faithfulness. The realization of the third person plural in the perfective aspect as the suffix {-u}, as exemplified in (15), reflects a systematic pattern

governed by positional constraints. This pattern underscores the consistent application of positional faithfulness principles across various phonological processes in MA. By analyzing such phenomena through the lens of positional faithfulness, scholars can gain deeper insights into the underlying phonological mechanisms that shape linguistic structures and forms.

(15)

xərʒ-u	'they left'
ʃəTħ-u	'they danced'
baʕ-u	'they sold'
ləʕb-u	'they played'

The phenomenon of glide formation in MA presents a compelling case for applying positional faithfulness constraints in phonological analysis. Specifically, when the root ends in a vowel, the suffix {-u} undergoes a transformation into the glide /w/. This process resolves hiatus structures, a phonological challenge in MA, by introducing a consonantal element to improve syllable structure and satisfy the markedness constraint ONSET, which prohibits onsetless syllables (V). Interestingly, positional faithfulness predicts that the affix, rather than the root vowel, undergoes modification in this process. This observation highlights the systematic nature of positional faithfulness in governing phonological repairs and optimizing syllable structures in MA.

(16)

ʒa-w	**ʒa-u*	'they came'
mʃa-w	**mʃa-u*	'they went'
bka-w	**bka-u*	'they cried'
ʕTa-w	**ʕTa-u*	'they gave'
ʃka-w	**ʃka-u*	'they complained'

In view of these facts, Noamane (2020) extended the framework of positional faithfulness to elucidate the morphological causatives in MA. Central to his proposal is the notion that the first radical consonant of the root occupies a privileged position in MA, resisting phonological and morphological processes. To formalize this observation, Noamane introduces the positional constraint IDENT-RtC1 (Weight), where "Rt" represents the root and "C1" denotes the root-initial consonant[6]. This constraint asserts that the

[6] While Noamane's proposal introduces a constraint specifically tailored to the Semitic linguistic context, we find it necessary to critically assess its universality within the OT framework. Our contention arises from the observation that certain languages, such as English, do not adhere to the notion of word formation revolving around a central root, which is commonly conceived in Semitic languages. Instead, Beckman's (2004) constraint IDENT-σ1 appears to capture a similar generalization with broader applicability across languages, suggesting a more universal perspective. Despite this, for the sake of exposition and consistency with prior analyses, we opt to adopt Noamane's constraint, albeit acknowledging its limitations in capturing cross-linguistic phenomena.

underlying feature specifications of the root-initial consonant must remain unchanged. The proposed constraints and their ranking are as follows:

(17)
a. *Constraints:*
RM: Some phonological exponent must appear in the output form.
IDENT-IO (Weight): Output segments and Input segments must be featurally identical for weight.
ALIGN-(μ_c, Left, Root, Left): The left edge of the causative morpheme must coincide with the left edge of the root.
IDENT-RtC$_1$ (Weight): The featural specification for the weight of the root's first radical element must be preserved in the input/output mapping.

b. Ranking:
RM >> IDENT-IO (Weight) and IDENT-RtC$_1$ (Weight) >>ALIGN-L (μ_C, Rt)

Noamane's (2020) analysis of morphological causatives in MA faces two significant challenges, both theoretical and empirical. The theoretical issue revolves around the proposed Root-Affix Metaconstraint, which fails to align with Beckman's (1998) established ranking schema for positional asymmetries within OT. As shown in (13), a markedness constraint should typically intervene between privileged and non-privileged positions, with the former remaining undominated. However, Noamane's schema deviates from this principle because it lacks the expected configuration in which markedness constraints mediate between privileged and non-privileged positions (Noamane 2020).

Empirically, the Root-Affix Metaconstraint encounters problems when confronted with observed phonological phenomena in MA. Specifically, two phonological processes, namely secondary labial assimilation and long-distance consonant harmony, cast doubt on the assumption that morphological causatives strictly adhere to positional faithfulness effects. The data in (18) and (19) below are a case in point:

(18) *Labial assimilation*
a. UR

/xrʒ/	*xʷruʒ*	'go out'
/ʃrb/	*ʃʷrub*	'drink'
/ħrb/	*ħʷrub*	'run away'
/qtl/	*qʷtul*	'kill'
/ʃf/	*ʃʷuf*	'see'

b.

/DRb/	*DRəb*	'hit'
/ʃTħ/	*ʃTəħ*	'dance'
/ħyyd/	*ħəyyəd*	'move'
/ʒwb/	*ʒawəb*	'answer'
/ktb/	*Ktəb*	'write'

(19) *Long-distance consonant harmony*
a. UR

/zwaʒ/	*ʒwaʒ*	‘marriage’
/zlʒlan/	*ʒəlʒlan*	‘sesame seeds’
/zuʒ/	*ʒuʒ*	‘two’
/sRʒm/	*ʃəRʒəm*	‘window’
/sfnʒ/	*ʃfənʒ*	‘doughnut’
/zlliʒ/	*ʒəlliʒ*	‘tiles’

In the context of MA, both labial assimilation and consonant harmony phenomena challenge the positional faithfulness hypothesis. Labial assimilation entails the initial consonant undergoing labialization due to the influence of an adjacent high-round vowel when present in the root. Similarly, consonant harmony involves the regressive assimilation of the palatal consonant, resulting in a change of place for the initial consonant to post-alveolar positions (Zellou 2010). Intriguingly, in both cases, the initial consonant undergoes modification, contrary to the predictions of positional faithfulness. These observations suggest that the first radical consonant of the root does not occupy a privileged position, undermining the core premise of positional faithfulness.

Another criticism arises concerning the postulation of the “Realize Morpheme” (RM) constraint, raising doubts about its theoretical necessity and empirical efficacy (van Oostendorp 2005; Trommer 2008). Critics argue that its inclusion may not significantly enhance explanatory power beyond existing constraints like DEP-IO. Given that DEP-IO already addresses the presence of morphemes in the input, the addition of RM appears redundant in many instances. Moreover, the RM constraint lacks clear criteria for distinguishing between phonologically realized and unrealized morphemes, leading to ambiguity and inconsistency in its application. This ambiguity raises concerns regarding the constraint’s universality and generalizability, potentially violating typological principles by rigidly requiring the full phonological realization of morphemes (van Oostendorp 2005).

To recapitulate, previous analyses of geminated causatives in MA have explored templatic frameworks and positional faithfulness hypotheses. However, these approaches encounter limitations. The templatic analysis alone failed to adequately capture morphological causatives, whereas empirical evidence undermines the validity of positional faithfulness. Given these shortcomings, this paper shifts its focus toward presenting an alternative analysis. We will delve into the specifics of MA’s syllable structure, arguing that it serves as the primary driver for the infixal process, ultimately leading to the observed phenomenon of TETU.

6. Proposed analysis

TETU, an important phenomenon within OT (McCarthy & Prince 1994; Alderete et al. 1999), manifests when a constraint typically inactive in a language becomes active in specific contexts where higher-ranked constraints fail to select the optimal candidate. This

phenomenon is often observed in reduplication, where reduplicants prefer unmarked structures, even if they violate the language's phonotactics (Kennedy 2008). The ranking schema for reduplicative TETU, as proposed by Alderete et al. (1999), highlights the complex interaction between faithfulness and markedness constraints. Specifically, it demonstrates the dominance of $Faith_{IO}$ over markedness constraints, while markedness constraints, in turn, dominate $Faith_{BR}$. The ranking schema that captures this fact is given in (20):

(20)
Ranking schema for reduplicative TETU (Alderete *et al.* 1999: 330):
$Faith_{IO}$ >> M >> $Faith_{BR}$

Drawing parallels with reduplication, we argue that the causative morpheme, denoted as AffixCAUSEμ, exhibits a similar unmarked preference, leading to its emergence as an infix. This emergence is driven by the otherwise inactive constraint *$COMPLEX_{ONSET}$. Assuming that the causative morpheme functions as a consonantal mora aligns with the cross-linguistic observation that geminates are underlyingly moraic (Davis 1999; Davis & Torretta 1998; Noamane 2019). The proposed constraints for deriving causatives and their ranking are crucial for understanding this process.

(21)
a. *Constraints:*
ALIGN-Root-C2: Align the second consonant of the root with the corresponding position in the output.

ALING ($Affix_{CAUSE}$, L, Root, L): Align the left edge of the $Affix_{CAUSE}$ with theleft edge of the root = every $Affix_{CAUSE}$ is a prefix in the Root.
***Min-σ:** Minor syllables are prohibited.
MAX_{IO}: Every segment of the input has a correspondent in the output.
DEP_{IO}: Every segment of the input has a correspondent in the output.
***$COMPLEX_{CODA}$**: More than one consonant in the coda position is prohibited.
***$COMPLEX_{ONSET}$**: More than one consonant in the onset position is prohibited.

b. *Ranking:*
MAX_{IO}, $COMPLEX_{CODA}$, *Min-σ >> DEP_{IO} >> *$COMPLEX_{ONS}$ >> ALING- $Affix_{CAUSE}$μ-L.

The constraints proposed for the analysis of causative formation in this framework offer a comprehensive account of the phonological and morphological processes involved. ALIGN-Root-C2 plays a crucial role in ensuring that the second consonant of the root aligns properly in the output, reflecting the typical position where causative morphemes are inserted. This constraint captures a fundamental aspect of causative formation and contributes to the systematicity of morphological processes. Similarly, ALIGN (Affix-CAUSE, L, Root, L) highlights the consistent morphosyntactic behavior observed in causative constructions, where the causative affix typically appears as a prefix within the

root. By enforcing this alignment, the constraint reflects cross-linguistic tendencies in morpheme positioning.

The *Min-σ constraint serves to maintain the phonotactic well-formedness of derived forms by prohibiting the formation of minor syllables, which contain fewer than two moras. This ensures that syllable structures adhere to the phonological constraints of the language and contribute to the overall prosodic harmony of the derived causatives. In addition, MAX_{IO} ensures that every segment present in the input is preserved in the output, preventing any deletion of phonetic material. DEP_{IO}, on the other hand, prevents the addition of new segments in the output that were not present in the input, thereby prohibiting epenthesis. Together, these constraints maintain both phonological transparency and morphological integrity in the derivation of causatives, ensuring that the output accurately reflects the phonological and morphological properties of the input roots.

In terms of ranking, MAX_{IO} and *Min-σ are given higher priority, reflecting the importance of preserving input segments and adhering to phonotactic constraints. DEP_{IO} follows, emphasizing the need for faithful preservation of input segments. Finally, *$COMPLEX_{CODA}$ and *$COMPLEX_{ONSET}$ constrain the complexity of coda and onset consonant clusters, respectively, ensuring that the derived forms adhere to the language's phonotactic patterns. ALING-AffixCAUSE-L is positioned last, reflecting its role in maintaining the morphosyntactic structure of causative constructions. Overall, these constraints provide a comprehensive framework for analyzing causative formation, capturing both phonological and morphological aspects of the process. An example of how these constraints interact is given in (22), where candidate (c) incurs the least violation; hence, it is selected as optimal:

(22)

Input : / $Affix_{CAUSE}$μ-ktb/	MAX_{IO}	*$COMP_{CODA}$	*Min-σ	DEP_{IO}	*$COMP_{ONS}$	ALING
a. ktb			*			
b. kkətəb				*!*	*	
☞ c. kəttəb				*!*		*
d. ktəb.b			*	*		
e. ktət.b			*	*		*
f. ktəbb		*		*	*	*
g. ktətb		*		*	*	*

In the evaluation of candidate forms for causative formation, constraints play a crucial role in determining the optimal output. Candidate (a) is immediately ruled out because it violates the higher-ranked constraint, which requires the realization of the AffixCAUSEμ. This underscores the importance of prioritizing constraints that ensure the faithful realization of morphological elements in the derived forms. Similarly, candidates (d) and (e) are eliminated due to their violation of *Min-σ, which prohibits the formation of minor syllables. This constraint reflects a phonotactic requirement to maintain the structural integrity of syllables in the language.

Candidates (f) and (g) face disqualification for violating the top-ranked constraint *COMPLEX$_{CODA}$, which restricts the presence of multiple consonants in coda positions. This constraint reflects a phonological restriction on syllable structures, ensuring that the derived forms adhere to the language's phonotactic patterns. The remaining candidates demonstrate a tie in violation of the active constraints, highlighting the need for a mechanism to resolve such conflicts.

Of particular interest is the tie between the two remaining candidates regarding their violation of the faithfulness constraint DEP$_{IO}$. Both candidates incur the same number of violations of this constraint, which emphasizes the importance of maintaining faithfulness to the input segments in the derived forms. Excluding the constraints where no violations occur, the evaluation process narrows down to prioritize constraints that ensure phonotactic well-formedness and faithful preservation of input segments, leading to the selection of the optimal candidate.

(23)

Input: /Affix$_{CAUSE}$μ-ktb/	DEP$_{IO}$	*COMP$_{ONS}$	ALING
b. kkətəb	*!*	*	
☞ c. kət.təb	*!*		*

In the final stage of candidate evaluation, the resolution of ties is crucial for determining the optimal output. Adopting Prince and Smolensky's *Method of Mark Cancellation* provides a systematic approach to address such ties by canceling out the same number of violation marks shared by competing candidates. In the case at hand, candidates (d) and (c) violate the faithfulness constraint DEP$_{IO}$, leading to their cancellation. Consequently, candidate (b) is ruled out due to its violation of *COMPLEX$_{ONSET}$, despite the irrelevance of this constraint in the language's phonological context. This outcome underscores the operation of TETU, where otherwise inactive constraints become active under specific conditions, thus shaping the derivation of morphological forms.

The emergence of morphologically derived causatives in MA as instances of TETU offers insights into the interaction between morphological and phonological constraints. The fact that *COMPLEX$_{ONSET}$ dominates the alignment constraint highlights the influence of phonotactic considerations on the morphological process. This observation aligns with cross-linguistic findings suggesting that dominant prosodic demands may override morphological requirements, as evidenced by the data on causative formation.

Despite the selection of the optimal candidate, potential forms such as *kəbtəb and *kəktəb appear to comply with the language's phonotactic constraints. However, these candidates are ruled out by invoking the constraint NO-CROSSING, which prohibits the long-distance spreading of consonants. This constraint serves to maintain the structural integrity of syllables and prevents violations of phonotactic patterns. Thus, the NO-CROSSING analysis provides a principled account for the exclusion of these potential candidates, further bolstering the analysis of morphological causatives in MA.

(24)

Since these two candidates do not surface, this constraint must be higher ranked, as shown in the following tableau:

(25)

Input: /RED-ktb/	NO-CROSSING	MAX_{IO}	$*COMP_{CODA}$	*Min-σ	DEP_{IO}	$*COMP_{ONS}$	ALING
a. ktb				*			
b. kkətəb					* !*		
☞ c. kəttəb					* !*		*
d. ktəb.b				*	*		
e. ktət.b				*	*		*
f. ktəbb			*		*	*	*
g. ktətb			*		*	*	*
h. kəbtəb	* !						
i. kəktəb	* !						

The strength of the present analysis lies in its ability to provide a unified and theoretically grounded account of morphological causatives in MA. Unlike the templatic analysis and positional faithfulness approaches, which have been previously proposed but have encountered theoretical and empirical challenges, the analysis presented in this paper offers a more robust and empirically supported framework.

First, the present analysis, which is based on OT, allows for a more principled explanation of morphological causatives in MA by integrating both phonological and morphological constraints. By using constraints such as ALIGN-Root-C2, ALING (AffixCAUSE, L, Root, L), *Min-σ, MAX_{IO}, DEP_{IO}, $*COMPLEX_{CODA}$, and $*COMPLEX_{ONSET}$, the analysis captures the intricate interplay between phonological and morphological factors in the formation of causative verbs.

Second, the present analysis demonstrates its superiority over templatic approaches by moving away from the rigid template-based morphology and instead adopting a constraint-based approach that can account for both regular and exceptional patterns observed in MA. Unlike templatic analyses, which often struggle to accommodate exceptions or irregularities, the OT framework allows for a more flexible and adaptive analysis that capture the full range of data without resorting to ad hoc stipulations.

Similarly, the analysis presented in this paper surpasses the positional faithfulness approach by offering a more nuanced and empirically supported account of morphological causatives in MA. While positional faithfulness has been proposed to explain certain phonological patterns in morphological processes, its application to MA causatives faces challenges due to the presence of other phonological processes that seem to override

positional constraints. The OT-based analysis, on the other hand, can accommodate these competing constraints and provide a more comprehensive explanation of the data.

Equally importantly, our analysis of TETU in MA causatives has broader implications for understanding other linguistic phenomena. For instance, the tendency for reduplicated forms to exhibit unmarked structures aligns with our findings. Reduplication, a prevalent morphological process in many languages, often involves the repetition of a morpheme or part of a morpheme. In many cases, the reduplicant appears with an unmarked structure, mirroring the patterns observed in MA causatives (McCarthy & Prince 1994). This parallelism suggests that the Emergence of the Unmarked may be a more widespread phenomenon, with implications beyond the specific context of MA causatives (Pecker & Potts 2011).

7. Conclusion

In conclusion, this study has undertaken a comprehensive investigation into the derivation of morphologically derived causatives in MA, employing OT as the theoretical framework. Through a critical examination of previous analyses, including templatic approaches and positional faithfulness accounts, we identified theoretical and empirical shortcomings that underscore the need for a more nuanced explanation. Building upon these insights, our analysis offers a novel perspective that integrates phonological and morphological constraints to elucidate the process of infixation in MA causatives.

The application of TETU has emerged as a central tenet in our analysis, shedding light on the dynamic interaction between phonological and morphological constraints in language. By demonstrating how otherwise inactive constraints can become operational under specific conditions, we highlighted the intricate interplay between phonotactic considerations and morphological processes. This has led to a deeper understanding of how dominant phonological demands may override morphological requirements, as evidenced by the emergence of infixation in MA causatives.

Moreover, our analysis addressed empirical challenges posed by phonological processes such as labial assimilation and consonant harmony, which question the validity of positional faithfulness effects in MA. By highlighting how these phenomena are better accounted for within the framework of TETU, we underscore the robustness and explanatory power of our proposed approach.

In summary, our study advances our understanding of morphological causatives in MA by offering a principled analysis grounded in OT principles. By integrating phonological and morphological constraints within a unified framework, we provide a more nuanced and empirically grounded account of infixation processes in MA. This not only contributes to our theoretical understanding of language structure and derivation but also underscores the importance of considering the intricate interplay between phonology and morphology in linguistic analyses.

Abbreviations

M	Mora	IO	Input-Output
Σ	Syllable	MAX	Maximality
☞	Optimal candidate	Min	Minor
*	Violation mark	Ms	Masculine
*!	Fatal violation	OT	Optimality Theory
3	Third person	RM	Realize Morpheme
ALIGN-L	Left Alignment	RtC_1	First radical segment
C	Consonant	SA	Standard Arabic
DEP	Dependency	TETU	The Emergence of the Unmarked
F	Faithfulness	UR	Underlying Representation
IDENT	Identity	V	Vowel
MA	Moroccan Arabic		

Conflict of Interest: The author declares no conflict of interest
Authors' Contributions: The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

References

Al Ghadi, Abdellatif. 2014. *Moroccan Arabic Plurals and the organization of the lexicon*. Rabat : Publications of the Faculté des Lettres et des Sciences Humaines de Rabat.

Al Ghadi, Abdellatif. 2022. An OT account of Moroccan Arabic prosody. *International Journal of Arabic Linguistics* 8, 16-32.

Alderete, John. 2001. Root-controlled accent in Cupeño. *Natural Language and Linguistic Theory* 19, 455-502.

Alderete, John & Beckman, Jill & Benua, Laura & Gnanadesikan, Amalia & McCarthy, John & Urbanczyk, Suzanne. 1999. Reduplication with fixed segmentism. Linguistic Inquiry 30, 327-364.

Bahloul, Maher. 2008. *Structure and function of the Arabic verb*. London: Routledge.

Beckman, Jill. 1998. *Positional faithfulness*. University of Massachusetts, Amherst. (Doctoral dissertation.)

Beckman, Jill. 2004. Positional faithfulness. In McCarthy, John (ed.), *Optimality Theory: A reader*, 310-342. Oxford: Blackwell.

Benhallam, Adberrafi. 1990. Moroccan Arabic syllable structure. *Langues et littératures* 8, 177-191.

Benhallam, Adberrafi & Dahbi, Mohamed. 1990. Accents of Moroccan Arabic: A preliminary study. In Pleines, Jochen (ed.), *La linguistique au Maghreb/Maghreb Linguistics*, 111-125. Rabat: Okad.

Benmamoun, Abbas. 1991. Causatives in Moroccan Arabic. In Comrie, Bernard & Eid, Mushira (eds.), *Perspectives on Arabic linguistics III*, 173-195. Amsterdam: John Benjamins.

Bennis, Said. 1992. *La formation du causatif en arabe marocain*. Rabat: Mohammed V University, Faculty of Letters. (D.E.S. thesis).

Bennis, Said & Iazzi, El Mehdi. 1995. Morphologie du verbe causatif en arabe marocain. *Langues et littératures* 13, 79-96.

Bensoukas, Karim, & Boudlal, Abdelaziz. 2012. The prosody of Moroccan Amazigh and Moroccan Arabic: Similarities in the phonology of schwa. In Borowsky, Toni & Kawahara, Shigeto & Sugahara, Mariko & Shinya, Takahito (eds.), *Prosody matters: Essays in honor of Lisa Selkirk*, 3-42. London: Equinox.

Boudlal, Abdelaziz. 2001. C*onstraint interaction in the phonology and morphology of Casablanca Moroccan Arabic*. Rabat: Mohammed V University, Faculty of Letters. (Doctoral dissertation.)

Davis, Stuart. 1999. On the moraic representation of underlying geminates: Evidence from prosodic morphology. In Kager, René & van der Hulst, Harry & Zonneveld, Wim (eds.), *The prosody morphology interface*, 39-61. Cambridge: Cambridge University Press.

Davis, Stuart & Torretta, Gina. 1998. An Optimality-Theoretic account of compensatory lengthening and geminate throwback in Trukese. *Papers from the Annual Meeting of the North East Linguistic Society* 28, 111-125.
Goldsmith, John. 1976. *Autosegmental phonology*. Cambridge, MA: MIT Press. (Doctoral dissertation.)
Harrell, Richard S. 1962. A *short reference grammar of Moroccan Arabic*. Washington, D.C.: Georgetown University Press.
Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
Kennedy, Robert. 2008. Bugotu and Cheke Holo reduplication: In defence of the Emergence of the Unmarked. *Phonology* 25(1), 61-82.
Lombardi, Linda. 2004. Positional faithfulness and voicing assimilation in Optimality Theory. In McCarthy, John (ed.), *Optimality Theory: A reader*, 343-364. Oxford: Blackwell.
Loutfi, Ayoub. 2014. *The morphosyntax of verbs and argument selection in Arabic*. Saarbrücken: Lambert Academic Publishing.
Loutfi, Ayoub. 2017. A*spects of the clause structure and word formation in Arabic: A Distributed Morphology analysis*. Rabat: Mohammed V University. (Doctoral dissertation.)
Loutfi, Ayoub. 2020. Deriving morphological causatives in Moroccan Arabic. *Macrolinguistics* 8(1), 1-17.
Loutfi, Ayoub. 2024a. The distribution of verbs and argument selection in Arabic: A Distributed Morphology account. In Jaḥfa, ʕAbd al-Majīd & Aġlīmū, Muḥammad (eds.), *Al-fiʕl fī al-ʕarabiyya: Bayn al-ṣarf wa-l-dalāla wa-l-tarkīb*, 223-290. Casablanca: Faculty of Letters and Human Sciences, Ben M'sik.
Loutfi, Ayoub. 2024b. *Causatives in Arabic: A morphosyntactic analysis*. Unpublished manuscript, Hassan II University of Casablanca, Morocco.
McCarthy, John. 1979. *Formal issues in Semitic phonology and morphology*. Cambrdighe, MA: MIT. (Doctoral dissertation.)
McCarthy, John. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373-418.
McCarthy, John. 1993. Templatic form in prosodic morphology. *Proceedings of the Formal Linguistics Society of Mid-America* 3, 187-218.
McCarthy, John, & Prince, Alan. 1994. The Emergence of the Unmarked: Optimality in prosodic morphology. *North East Linguistic Society* 24, 333-379.
McCarthy, John, & Prince, Alan. 1999. Faithfulness and identity in prosodic morphology. In Kager, René & van der Hulst, Harry & Zonneveld, Wim (eds.), *The prosody morphology interface*, 218-309. Cambridge: Cambridge University Press.
McCarthy, John, & Prince, Alan. 2004. Generalized alignment: The prosody-morphology interface. In McCarthy, John (ed.), *Optimality theory in phonology: A reader*, 451-463. Oxford: Blackwell.
Noamane, Ayoub. 2019. The moraic nature of geminate consonants in Moroccan Arabic: Evidence from word-minimality, syllable structure and word formation. *International Journal of Arabic Linguistics* 5(2), 100-129.
Noamane, Ayoub. 2020. The morphophonology of Moroccan Arabic derived causatives. *Macrolinguistics* 8, 1-27.
Prince, Alan & Smolensky, Prince. 2004. *Optimality Theory: Constraint interaction in Generative Grammar.* Malden, MA, & Oxford: Blackwell.
Trommer, Jochen. 2008. Multiple-feature mutation and Realize Morpheme. *Linguistische Arbeitsberichte* 87, 163-182.
Urbanczyk, Suzanne. 2007. Reduplication. In De Lacy, Paul (ed.), *The Cambridge handbook of phonology*, 473-493. Cambridge: Cambridge University Press.
Ussishkin, Adam. 2000. The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology* 16, 401-442.
Ussishkin, Adam. 2007. Morpheme position. In De Lacy, Paul (ed.), *The Cambridge handbook of phonology*, 457-472. Cambridge: Cambridge University Press.
van Oostendorp, Marc. 2005. Expressing inflection tonally. *Catalan Journal of Linguistics* 4, 107-127.
Yip, Moira. 1991. Coronals, coronal clusters and the coda condition. In Paradis, Carole & Prunet, Jean-Fancois (eds.), *Phonetics and phonology 2: The special status of coronals*, 61-78. New York: Academic Press.
Yu, Alan C. L. 2007. *A natural history of infixation*. Oxford: Oxford University Press.
Zellou, Georgia. 2010. Moroccan Arabic consonant harmony: A multiple causation hypothesis. *Toronto Working Papers in Linguistics* 33. (https://twpl.library.utoronto.ca/index.php/twpl/article/view/6780) (Accessed 2024-06-21.)

DOI: 10.14746/linpo.2024.66.1.3

Delineating a case system: How many cases are there in Standard Goan Konkani – and why?[1]

John Peterson[1] & Govind Mopkar[2]

[1] Kiel University
jpeterson@isfas.uni-kiel.de | ORCID: 0009-0007-7941-3666
[2] Deccan College, Post Graduate and Research Institute, Pune
mopkar.govind.n@gmail.com | ORCID: 0009-0005-9821-4561

Abstract

In the present study we analyze the case system of Standard Goan Konkani, an Indo-Aryan language spoken in western India. We first analyze the individual markers which fit our definition of case markers and discuss how these differ from and/or overlap functionally with other case markers. We then present a detailed description of two further case markers which, although quite common, to our knowledge have not been discussed in previous works on Konkani, namely the *selective* and the *elative*, and show how these forms, despite their transparent etymologies as case stacking of the inessive case and either the genitive or the ablative, cannot be viewed as case stacking in the modern language. With this, we argue that Konkani can best be described as having 13 productive cases, although other analyses are possible, depending on researchers' theoretical and practical preferences.

Keywords: Konkani, Indo-Aryan, case system, selective case, elative.

1. Introduction

The present study deals with the case system of Standard Goan Konkani (ISO 639-3: gom; Glottocode: goan1235), an Indo-Aryan language spoken primarily in western India.

[1] We would like to thank Jussi Ylikoski and Andrea Santamaria for their invaluable comments and suggestions on an earlier version of this article. Needless to say, of course, that all remaining inconsistencies and errors are the sole responsibility of the authors. The first author would also like to thank the German Research Council (DFG) for funding the project "Towards a linguistic prehistory of eastern-central South Asia (and beyond)" (Project no. 326697274), which made field work for this study possible.

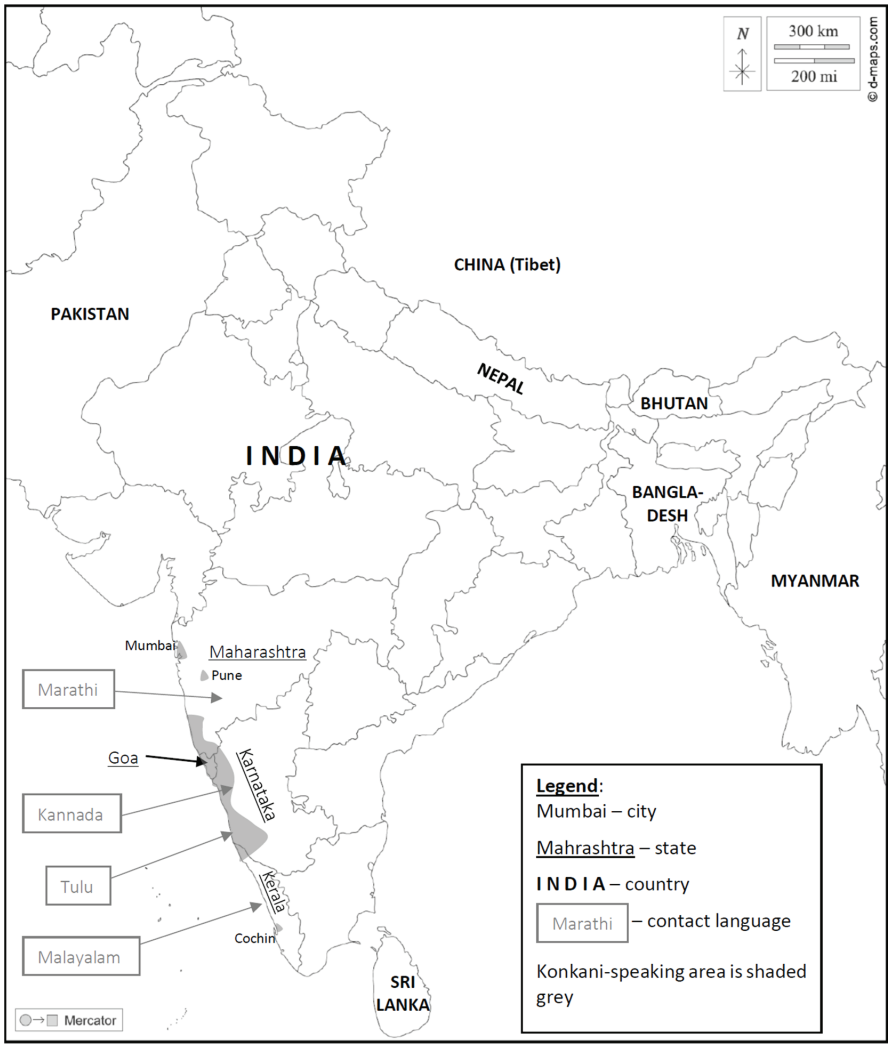
Konkani ([koŋkɳiː]) is the official language of the state of Goa, the only region where it is spoken by a majority of the population. Outside of Goa Konkani is spoken in different varieties as a minority language throughout a relatively narrow strip of land along the west coast from the state of Maharashtra in the north, through Goa, and in much of coastal Karnataka to the south. There are also small pockets of Konkani in and near Pune and Mumbai in Maharashtra, and in Cochin in the southwestern state of Kerala (cf. Almeida 1989: 5-7). Konkani is thus in close contact with the Indo-Aryan language Marathi in Maharashtra, with the Dravidian languages Kannada and Tulu in Karnataka, and with Malayalam (also Dravidian) in Kerala; see Map 1, based on Almeida (1989: 6). According to the *Ethnologue* (Eberhard et al. 2023), Goan Konkani was spoken in 2000 by 3,630,000 people in India and 3,725,300 in all countries.

Konkani is a "macro-language", defined by Eberhard et al. (2023) as "multiple, closely related individual languages that are deemed in some usage contexts to be a single language."[2] The different varieties of Konkani throughout the region where these are spoken can thus show considerable differences. In this situation, a de facto standard of Konkani has emerged which we refer to here as "Standard Goan Konkani" and which is based on the Goan Hindu dialect. This "standard" is not to be understood in the sense of one single "correct" form of speech which has been standardized by an organization with the recognized authority to determine which forms of the language are or are not standard. To our knowledge, this type of standard does not exist for any form of Konkani. Rather, with this term we refer to the socially recognized standard form of Goan Konkani for which there is a stable, agreed upon orthography, and for which there are also morphological categories and syntactic constructions which are recognized by all educated speakers and which are expected in written and in formal spoken language.

Despite this generally agreed upon written standard, as we show in the following pages there is presently no real consensus among researchers with respect to the case system of this language – with the number of posited cases ranging from six to eleven in individual studies – and it is not always clear how the respective authors have reached the number of cases that they assume. With the present study we hope to shed some light on the problems and issues involved. While we do not claim to have found the "correct" interpretation of the Konkani case system, we will suggest our own analysis and show how this differs from other analyses – and why.

The remainder of this article is structured as follows. In Section 2 we give the definition of *case* that we use in this paper and present our criteria for differentiating between case markers and postpositions in Konkani. In Section 3 we give an overview of the case system as this has been analyzed in a number of previous studies of Konkani, stretching back some 100 years. In the following two sections we discuss various forms in considerable detail and the issues involved in determining their status as separate cases: In Section 4 we discuss various genitive case forms, based on the results of a small corpus analysis with respect to claims made in the literature, and in Section 5 we take a closer look at the various markers of the ablative and related forms.

[2] https://www.ethnologue.com/methodology/ (Last accessed 2023-07-14).



Map 1: The Konkani-speaking regions and Konkani's major contact languages, based on Almeida (1989: 6)[3]

In Section 6 we then present evidence for two further cases in Konkani, hitherto unmentioned in the literature, namely the selective and the elative. In Section 7 we then return to our main question, namely how many cases there are in Konkani – and why, taking the uncertainties into account that we discuss in this article. Finally, Section 8 presents a summary of this study with a brief outlook for future research.

[3] This map is a modified version of the map given at https://d-maps.com/carte.php?num_car=4183&lang=en, to which the names of states, countries, the Konkani language and the legend have been added. (Last accessed: 2024-05-23.)

2. Case markers and postpositions

We follow here the definition of case given in Blake (1994: 1): "**Case** is a system of marking dependent nouns for the type of relationship they bear to their heads. Traditionally the term refers to inflectional marking, and, typically, case marks the relationship of a noun to a verb at the clause level or of a noun to a preposition, postposition or another noun at the phrase level." (emphasis in the original) In addition to cases defined in this way and following Blake (1994: 9), we also assume that the vocative, i.e., the form of a noun used to address someone, can be considered a case for structural reasons if it forms part of the same nominal paradigm to which the other cases belong, as is the case, e.g., in Latin and Greek.

Case in Konkani is not predominantly marked inflectionally but rather enclitically, by markers which likely derive from, but are also distinct from, postpositions. Miranda (2003: 744) assigns case markers in Konkani to the larger class of postpositions, which he divides into two groups. The first group, consisting of case markers, are what Miranda refers to as "clitic postpositions", the second group being "secondary postpositions". With respect to this distinction Miranda (2003: 744) writes:

> With the exception of the superessive, clitic postpositions occur after the oblique form of a noun or pronoun.[4] Secondary postpositions can occur after the oblique form of a noun or non-personal pronoun or after the genitive oblique singular form of the same […], but they must occur only after the genitive oblique singular form of a personal pronoun […].

Dhongde (2022: 53-54) notes that "[a] clitic may intervene between a postposition and the governing noun", e.g. in (1), but that this is not possible with case markers, as shown in (2).[5] Note that the "clitic" *=ʧya* in both examples is in fact the genitive case marker, so that Dhongde's argumentation also supports Miranda's analysis.

(1)	*ram-a=pasun* /	*ram-a=ʧya=pasun*	*ʧiʈ*	*ghe.*	
	Ram-OBL=PP	Ram-OBL=clitic=PP	letter	take.IMP	
	'Take the letter from Ram.'				
(2)	*ram-a=k*	/	**ram-a=ʧya=k*	*ʧiʈ*	*di.*[6]
	Ram-OBL=DAT	/	Ram-OBL=clitic=DAT	letter	give.IMP
	'Give the letter to Ram.'				

[4] Unlike other case markers, the superessive superessive marker *=r* is only added to personal pronouns marked by the genitive *=ce*. Cf. Section 2 for further discussion.

[5] We have silently corrected a small printing error here from Dhongde (2022: 54). Also, we indicate clitics with "=" and suffixes with "-" in Dhongde's examples and gloss according to the Leipzig Glossing Rules, where possible, in keeping with the rest of the present article.

Dhongde (2022: 53) refers to both Miranda's (2003) class of postpositions and the direct and oblique stems of nouns as "cases". In contrast, we refer in the present article to these latter forms as the direct and oblique stems and reserve the term "case" for Miranda's (2003) "clitic postpositions". Cf. §§3.1 and 3.2 for further discussion.

[6] The second form in (1), with *ram-a=ʧya=pasun*, appears to be from a non-Goan variety. Furthermore, (2) with *ram-a=ʧya=k*, is in fact grammatical if this is viewed as case stacking of the genitive *=ʧya* and the "dative" *=k*, but not with the meaning intended in (2). Instead it would mean 'Give the letter to Ram's [child, etc.]'. As this second meaning is irrelevant for our discussion here we do not discuss this issue further.

We follow the argumentation in Miranda (2003), although we will speak of "postpositions" for those forms which must take the genitive with pronouns vs. "case markers", which do not require the genitive with pronouns. There are however two exceptions to this second group:

- The first is that we consider units which transparently derive from the genitive plus a case marker, but which are entirely interchangeable with forms lacking this genitive marker in all positions outside of the pronominal system, to be case markers. For example, the superessive marker *=cer*, which consists of the genitive marker *=c* marked for the oblique singular marker by *-e*[7] plus the superessive marker *=r*, is interchangeable with the simple marker *=r* in all environments outside of the pronominal system. We therefore view it as a simple variant of the form *=r*, which is unambiguously a case marker.
- Secondly, the marker *=cyan*, which derives from the genitive marker *=c* followed by the masculine / neuter singular marker *-ya* and the ergative / instrumental *=n*, is entirely interchangeable in its ablative function with all other ablative forms such as the free variants *=san*, *=sun*, *=than*, etc. (cf. Section 5) and is therefore a case marker in our view. This analysis is further supported by the fact that the ablative meaning of *=cyan* does not derive compositionally from the genitive and the ergative / instrumental case meanings, so that it has developed further as a case marker from only marking the ergative / instrumental to include the ablative.[8]

As the discussion of *=r* / *=cer* and *=cyan* above shows, when discussing the Konkani case system several choices with respect to the status of particular forms as case markers or as postpositions have to be made, and opinions will often differ among researchers. As the following pages illustrate, the cumulative effect of decisions such as these directly impact the number of cases assumed by a particular researcher.

3. The Konkani case system – structural overview and previous accounts

Despite Konkani's status as a scheduled language,[9] comparatively little descriptive work has been done on it in comparison with other scheduled languages, and much of the work which has been done on it is dedicated either to non-Goan Konkani varieties or to non-standard varieties of Goan Konkani. Furthermore, in those works which do deal with Goan Konkani, it is notable that these studies often come to very different conclu-

[7] *-e* here is a fossilized form found in some pronominal forms and with certain postpositions. Its more common form is *-ya*.

[8] In Section 5 we will show that *=cyan* not only marks ablative semantics but also instruments. For the sake of presentation, we do not deal with this further here.

[9] The term "scheduled languages" refers to the 22 languages presently listed in the Eighth Schedule to the Indian Constitution which enjoy a privileged status in education and administration, etc.

sions with respect to the case system, ranging from six to eleven cases in the different analyses.

The present section begins with a general typological overview of the case system of Standard Konkani in Section 3.1, describing its alignment patterns and the direct and oblique stems found with all nouns and pronouns, while Section 3.2 presents an overview of the different case systems suggested for Konkani in previous studies. Here, we include only those grammars which claim to describe either (Standard Hindu) Goan Konkani or simply "Konkani". We do not, however, include any varieties which specify in their title a local, non-standard or non-Goan variety, such as Christian Karnataka Konkani (Almeida 1989), the Christian Bardeshi dialect of North Goa (Almeida 2012), the Konkani of South Kanara (Ghatage 1963) and Kankon (Ghatage 1968), Kudali (Ghatage 1965), Kunabi (Ghatage 1966) or Gawdi Konkani (Ghatage 1972; Karapurkar 1968). A discussion of the different Konkani varieties in comparison with one another will be the subject of a later study.

Section 3.3 then gives a brief overview of the complex NP-internal agreement system of Konkani in order to facilitate the interpretation of the examples contained throughout the rest of the present study. Finally, Section 3.4 discusses the main functions of these cases, with examples.

3.1. General introduction

Standard Goan Konkani is a verb-final language with a split ergative-alignment system: The "transitive subject" (A) appears in the ergative in the simple past tense and in the perfect, while the "intransitive subject" (S) in these categories appears in the nominative. In all other finite verbal categories such as the present, future, and the past imperfective, S and A both appear in the nominative case.

The "direct object" (O)[10] can either appear in the zero-marked nominative or in the objective case. To our knowledge no study has yet researched O-marking in Konkani in detail, but the general principles seem to apply to Konkani as they do to most other South Asian languages, with a strong tendency for non-human and non-definite objects to appear in the unmarked nominative case and human and/or definite objects to appear in the objective case. S also shows variable marking, as it appears in the ergative with certain nonfinite forms, such as the future participle, and in the nominative elsewhere.

Simplifying somewhat, the verb agrees in terms of person, number and in some categories also in terms of gender with a nominative-case marked S or A or with the nominative-case marked O in transitive clauses in the past tense and in the perfect. If there is no nominative-case form with which the verb can agree, the verb usually shows default agreement, i.e., the 3rd person singular, neuter, although some speakers allow the verb to show agreement with the objective-marked noun, i.e., semantic marking. Which

[10] Strictly speaking, these comments apply to both O and T (i.e., the "direct object" in both mono- and ditransitive clauses), but as O and T behave identically with respect to case marking, and in order to simplify the discussion, we will speak of both of these argument types collectively as "O".

forms of Konkani allow this and under which conditions remains to be studied in more detail.

All nouns in Konkani have an inherent grammatical gender (masculine, feminine or neuter) and two different stems in both the singular and the plural, referred to here as the direct and oblique stems. The direct stem is the citation form and, since the nominative is zero-marked (see following section), the direct stem is always homophonous with the nominative case. The oblique stem is the stem to which case markers and postpositions attach. There are some 33 different nominal inflectional classes and subclasses with respect to the direct and oblique stems (cf. Peterson 2022). Table 1 presents a few examples of these different inflections (here for the singular only).

Table 1: Examples of direct stems, oblique stems, and oblique stems plus case markers in Konkani

Direct stem	Oblique stem (SG)	Oblique stem plus case marker
ʃar 'city (M)'	*ʃar-a*	*ʃar-a=k* 'to the city' (=*k* 'OBJective')
khuɖ 'room (F)'	*khuɖ-i*	*khuɖ-i=nt* 'in the room' (=*nt* 'INESS')
ʃet 'field (N)'	*ʃet-a*	*ʃet-a=nt* 'in the field'
ʃaɭa 'school (F)'	*ʃaɭ-e*	*ʃaɭ-e=k* 'to the school'

3.2. The Konkani case system – previous accounts

In this section we present an overview of some of the case systems which have been suggested by various researchers of Konkani over the past ca. 100 years. We stress that the following discussion is by no means exhaustive and only serves to show some of the different previous analyses of this system.

We begin with the system suggested by Dalgado (2022), written ca. 100 years ago.[11] Dalgado assumes eight cases for Konkani, however as one of these, the locative, shows two different forms with two different meanings and is given as two different subcategories by Dalgado, we count them here as two different cases, yielding nine cases altogether.[12] Table 2 summarizes the case system proposed by Dalgado, using current terminology, with Dalgado's terminology in parentheses where the two differ. As Table 2 shows, one potential problem with respect to Dalgado's analysis – one which will come up again in the following pages – is the fact that the "accusative" has two different forms, both of which are identical to another case, namely the nominative and the dative.[13]

[11] Dalgado (2022) is the translated English edition of the Portuguese-language manuscript by Rev. Mgs. Dr. Sebastião Dalgado, the hand-written original of which was apparently written between 1916-1922 and which is now on display in the Central Library of Goa (Dalgado 2022: 6-7, 12, 14).

[12] Dalgado (2022: 44-45) posits a third locative case, in *-ĩ*, which we do not include in Table 2 as it is not productive.

[13] Here and in the following tables, the nominative (and the "accusative" when this has the same form as the nominative) is formed by the direct stem plus the null morph. All other case markers, including the

Table 2: The Konkani cases system in Dalgado (2022)

	Singular	**Plural**
Nominative	*Ø*	
Accusative – formally identical with either the nominative or the dative	*Ø / =k*	
Ergative / instrumental (“instrumental”)	*=n / =cyan*	*=nĩ / =cyan*
Dative	*=k*	
Ablative	*=sun / =san / than*, etc.	
Genitive	*=cɔ / =ci / =cɛ̃*[14]	
Inessive (“Locative I”)	*=nt*	*=nt / =nĩ*
Superessive (“Locative II”)	*=r*	
Vocative	*Ø*	*=no / =nu*

In contrast, Miranda (2003) gives paradigms for pronouns with five cases but does not provide declensions for lexical nouns. However, he also mentions the “subessive” case marker (“inessive” in our terminology) in his discussion of clitic vs. secondary postpositions (Miranda 2003: 744) so that we include this case in Table 3, with six cases in total.

Table 3: The Konkani case system in Miranda (2003)

	Singular	**Plural**
Nominative	*Ø*	
Accusative / dative	*=k*	
Ergative / instrumental (“agentive”)	*=n*	*=ni*[15]
Genitive	*=cɔ / =ci / =cɛ̃* **or** *=l-ɔ / =l-i / =l-ɛ̃*	
Superessive	*=r*	
Inessive (“subessive”)	*=t̃*	

Almeida (2004: 48, 65-66, 78) assumes, in our interpretation of his discussion, a system of eleven cases, summarized in Table 4. Note that Almeida assumes both a dative and an accusative case which are formally identical. Almeida (2004: 66) also discusses a further form of the genitive, *=gɛl-*, referred to in the following text as the kinship genitive, with the meaning ‘belonging to the household of’. We therefore include this as a separate case in Table 4.

zero-marked vocative singular, attach to the oblique stem. For the sake of presentation, all case markers are given here as enclitics and the various transliteration systems have been unified for the sake of presentation.

[14] The genitive forms in Table 2 and all following tables agree with the following head noun with respect to gender, number and direct/oblique status of the stem. The full declension of these forms is given in Section 3.3 below (Table 7).

[15] We find no examples of the ergative plural given for nouns (only the somewhat different forms for pronouns, cf. Miranda 2003: 743), so we have filled in this slot of Table 3 with the modern standard form.

Table 4: The Konkani case system in Almeida (2004)

	Singular	**Plural**
Nominative (“direct”)	*Ø*	
Accusative	*=k*	
Dative (formally same as accusative)	*=k*	
Ergative / instrumental (“agent / instrumental”)	*=n*	*=ni*
Inessive (“Locative 1”)	*=t̃*	*=ni*
Superessive (“Locative 2”)	*=r / =cer*	
Kinship essive (“Locative 3”)	*=ger*	
Ablative	*=san / =sun*	
Genitive (“possessive”)	*=cɔ / =ci / =cɛ̃* **or** *=l-ɔ / =l-i / =l-ɛ̃*	
Kinship genitive	*=gɛl-ɔ / =gel-i / =gɛl-ɛ̃*	
Vocative	*Ø*	*=no*

In his recent grammar, Dhongde (2022: 53-59) assumes eight cases, summarized in Table 5 (without the forms from the Mangalore dialect). We include the vocative in this list, which Dhongde also views as a case but which he discusses separately from the other cases. Like Almeida, Dhongde (2022: 55) assumes both an accusative and a dative case, again although they are formally identical, although he generally glosses the two uniformly as DAT. Note also that he groups all three locative case forms mentioned by Almeida (i.e., inessive, superessive and kinship essive) together under “locative” and does not distinguish further between these forms or their functions.

Table 5: The Konkani case system in Dhongde (2022)

	Singular	**Plural**
Nominative	*Ø*	
Accusative	*=k*	
Dative (formally same as accusative)	*=k*	
Ergative	*=n / =ɳ*	*=ni / =ɳi*
Ablative	*pɵsun*	
Genitive	*=cɔ / =ci / =cɛ̃*	
Locative	*=t, =cer / =ger, =ã*	*=t, =cer*
Vocative	*=a / =o (M) / =e (F)*	*=o*

With eleven cases, Almeida’s (2004) system is the most detailed of any of the case systems above and also served as the basis for discussions in other works by the Peterson and his associates (e.g., Peterson 2022; Peterson & Chevallier 2022; Peterson & Mopkar 2021), although with one minor revision, namely that these latter works do not consider Almeida’s dative and accusative to be two separate cases. Instead, as the direct object (O) is identical in form to the nominative when it is non-human/indefinite and

identical in form to the indirect object (G) when it is human/definite, these studies consider the direct object to be either in the nominative case when it is unmarked, or in the objective case when it is marked by the objective marker *=k*. Table 6 provides an overview of this system, together with the most common forms of these markers.

Table 6: The modified system of Almeida (2004) as found in Peterson (2022), Peterson & Chevallier (2022) and Peterson & Mopkar (2021)

	Singular	**Plural**
Nominative (= direct stem)	*Ø*	
The following enclitic case markers attach to the oblique stem:		
Objective	*=k*	
Ergative / Instrumental	*=n*	*=ni*
Locative cases		
Inessive (≈ 'in')	*=nt (=n)*	*=ni*
Superessive (≈ 'on')	*=r / =cer*	
Kinship essive ('at the home of')	*=ger*	
Ablative	*=san / =sun / =savn / =than / =cyan*	
Genitive cases		
Genitive (general)	*=cɔ / =ci / =cɛ̃* **or** *=l-ɔ / =l-i / =l-ɛ̃*	
Kinship genitive ('of the household of')	*=gɛl-ɔ / =gel-i / =gɛl-ɛ̃*	
Vocative	*Ø*	*=no*

As noted in Peterson & Mopkar (2021: 37, fn. 10), while Standard Goan Konkani distinguishes between the ergative / instrumental singular *=n* and the inessive singular *=nt*, although they have identical forms in the plural, in the spoken language of many speakers, as well as in works written in colloquial style (e.g. Murkuɳḍe 2015, a book of stories for children), the inessive singular is spoken and written as *=n*, resulting in identical marking of the ergative / instrumental and the inessive in both singular and plural, so that these are best considered a single case for these speakers. However, as we are discussing here solely the standard dialect, where the distinction between these two cases is maintained in both singular and plural, we will continue here to distinguish these two cases in the following.

Thus, as this section shows, not only do the different researchers assume different case systems, interpreting exactly what each of these individual authors views as a separate case, as opposed to an allomorph of a more general case, itself often involves a certain amount of interpretation.

3.3. NP-internal agreement patterns

The genitive markers given in Tables 2-6 agree with the noun they modify in terms of gender, number and direct/oblique-stem status. This agreement pattern is identical with that found with declinable adjectives. The agreement system of the genitive with the following head noun is illustrated in Table 7. For the sake of comparison, Table 8 shows the corresponding forms of declinable adjectives.

Table 7: The forms of the genitive markers

Stem	**Singular**			**Plural**		
	M	F	N	M	F	N
Direct	*=cɔ / =lɔ*	*=ci / =li*	*=cɛ̃ / =lɛ̃*	*=cɛ / =lɛ*	*=cyo / =lyo*	*=cĩ / =lĩ*
Oblique	*=cya / =lya*	*=ce / =le*	*=cya / =lya*	*=cya / =lya*		

Table 8: The forms of the declinable adjectives in attributive use

Stem	**Singular**			**Plural**		
	M	F	N	M	F	N
Direct	*-ɔ*	*-i*	*-ɛ̃*	*-ɛ*	*-yo*	*-ĩ*
Oblique	*-ya*	*-e*	*-ya*	*-ya*		

Examples of genitive and adjectival agreement are given in (3)-(5). The oblique singular marker preceding the genitive enclitic marker in example (3), i.e., *-a* in *ɵdhyay-a=c-ɛ̃* ‘of the chapter’, is that of the respective inflectional class to which *ɵdhyay* ‘chapter’ belongs. The marker ***-ɛ̃*** following the genitive marker *=c* agrees in gender, number and stem status with the following possessum *rupantɵr* ‘version’, which is neuter, singular and appears in the direct stem.

(3) *ɵdhyay-a=c-ɛ̃* *rupantɵr*
chapter.M.SG-OBL.M.SG=GEN-DIR.N.SG version.DIR.N.SG
‘the version of the chapter’

In example (4), the NP in (3) is modified by the noun *khɵŋḍ* ‘section’, which is masculine, singular and – as it is followed by the genitive marker – also appears in the oblique stem. The final *-ya* in the form *khɵŋḍ-a=c-ya* shows agreement with the following masculine, singular, oblique-stem noun *ɵdhyay-a=c-ɛ̃*, from example (3).

(4) *khɵŋḍ-a=c-ya* *ɵdhyay-a=c-ɛ̃* *rupantɵr*
section.M.SG-OBL.M.SG=GEN-OBL.M.SG chapter.M.SG-OBL.M.SG=GEN-DIR.N.SG version.DIR.N.SG
‘the version of the chapter of the section’

Finally, example (5) shows the larger, attested NP from which (3) and (4) were taken. Here we find the NP-initial demonstrative *hya*, the masculine, singular, oblique form of

the proximal demonstrative, agreeing in terms of oblique status, gender and number with the oblique-stem, masculine singular noun *khəɳḍ* 'section'. The example also includes the adjective *pəyl-* 'first' in attributive function to the noun *ədhyay* 'chapter'. Here *pəyl-* is marked as masculine, singular oblique, i.e., *pəyl-ya*, as it agrees with the masculine, singular oblique noun *ədhyay* 'chapter'.

(5) *hya* *khəɳḍ-a=c-ya* *pəyl-ya* *ədhyay-a=c-ɛ̃*
this.OBL.M.SG section.M.SG-OBL.M.SG=GEN-OBL.M.SG first-M.OBL.SG chapter.M.SG-OBL.M.SG=GEN-DIR.N.SG

rupantər
version.DIR.N.SG
'the version of the first chapter of this section'

In order to simplify the examples and for the sake of readability, in the following pages oblique marking will simply be glossed as "OBL" for singular oblique forms and as "OBL.PL" with plural oblique nouns, only indicating gender when this serves to clarify the example. Similarly, the direct stem will not be glossed unless this serves to clarify the example.

3.4. The individual cases shown in Table 6 and their functions

We briefly summarize here the forms and functions of the ten cases in Table 6 from our own previous study (Peterson & Mopkar 2021) and provide an example for each case. The purpose of this is to give the reader a better idea of this system, on which all of the above-mentioned studies agree at least to some extent, before refining this system somewhat and discussing two further cases which have so far escaped the attention of researchers of Konkani.

- the zero-marked **nominative or direct case** is the case of intransitive subjects (S) of finite verbs in general and of transitive subjects (A) of finite verbs in the present tense, the future, and the past habitual. It is also found with non-human and indefinite O's, as with *aḍo* 'fencing' in example (6) (where the A is omitted as it is known from context), or *pətrə* 'letter' in (8).

(6) *aḍo=y=bi* *ke-ll-ɔ.*
fencing.NOM.M.SG=ADD=ADD do-PST.PERF-M.SG
'We also put in fencing (literally: '[We] did fencing also').' (CP: 35)[16]

- the **ergative / instrumental** marks the A in the finite past tense (e.g., *ain* 'mother' in (8)), and in the past and present perfect, and both S and A with non-finite forms (e.g., future participle). It also marks instruments and can have a causal interpretation, as shown in (7).

[16] References given as abbreviations between brackets, as here, are from spoken texts in our corpus.

(7) *[…] hẏa bhõẏ-a=n to=y kheɭ=na […]*
this.OBL fear-OBL=INST 3SG.M=ADD play=NEG.COP.PRS.3SG
'[…] through / out of this fear he also does not play […]' (Almeida 2004: 176)

- the **objective** is used to mark indirect objects (G), and definite/human direct objects (O). In (8) it marks G, *baba* 'father', the indirect object.

(8) *ai=n baba=k pətrə bərəy-l-ɛ̃.*
mother.OBL=ERG father.OBL=OBJ letter.NOM.N.SG write-PST-N.SG
'Mother wrote a letter to father.' (Almeida 2004: 121)

- the **inessive** marks a local relationship where one entity is contained within another, as in (9).

(9) *biɭ-a=nt ek muyed-ã=c-i poʈli asa.*
hole-OBL-INESS one coin-OBL.PL=GEN-F.SG bundle.F COP.PRS.3SG
'In the burrow there is a bundle of coins.' (Nayɔk 2017: 8)

- the **superessive** marks a local relationship in which one entity is physically on another; it is also used with reference to text/spoken language contained in different media, i.e., in a newspaper, on television or on the radio. (10) presents an example of this case in its primary meaning of 'on'.

(10) *mɛj-a=r dəvər-lã tɛ̃ pustək mhaka jay.*
table-OBL=SUPERESS place-PERF.3SG.N that.N.SG.DIR book.N.SG 1SG.OBL be.necessary
'I need the book that I placed on the table (lit.: 'I have placed on the table that book is necessary to me)'. (Almeida 2004: 173)

- the **kinship essive** expresses a locative relation with respect to a household or, less commonly, motion towards a household, referenced by one of the members of that household ('at the home of'). (11) provides an example of the first meaning.

(11) ... *tãɳi dilli lagĩ apəl-ya put-ã=ger rav-p-i ɖɔ*
3.HON.ERG Delhi close.to REFL-OBL.HON son-HON.OBL=KIN.ESS stay-NMLZR-ADJVZR Dr.

joʃĩ=c-ɔ namɔ ani fon nəmbər mhaka di-l-ɔ.
Joshi-OBL.HON=GEN-M.SG name. and. address.M andphone number 1SG.OBJ give-PST-M.SG
'… he gave me the name, address and phone number of Dr. Joshi, who was living at his son's home close to Delhi.' (Miranda 2019: x)

- the **ablative** refers to movement away from a source (physically or temporally), as in (12).

(12) *hãv karnaʈək-a=sun [ay-lã][17] haŋar kam kər-p-a=k.*
1SG Karnataka-OBL=ABL come-PERF.M.1SG here work do-NMLZR-OBL=OBJ
'I came here from Karnataka to work.' (RV: 4)

[17] Pronounced [ela] by this speaker.

- the **general genitive** is the default adnominal case, in which one NP is incorporated into another NP in attributive function, including but not restricted to possession. It can also be used to incorporate postpositions and adverbials into an NP as attributes, as in (13).

(13)

						pɵɭẽvk	
mhaka	*haŋa=c-yo*	*veɭ-o*	*ani*	*dharmik*	*jag-ɛ*	*pɵɭe-ũk*	*jay.*
1SG.OBL	here=GEN-F.PL	beach-F.PL	and	religious	place-M.PL	see-INF	be.wanted

'I would like to see the beaches and religious places of here.' (Almeida 2004: 59)

- the **kinship genitive** has the same function as the general genitive but is only used with a possessor NP which refers to the members of a household, referenced by one of the members of that household as possessor, as shown in (14).

(14) ... *mha=gɛl-ɔ* *ekʈ-ɔ* *frɛɳɖ*
1SG=KIN.GEN-M.SG one.single-M.SG friend.M.SG
'… a friend of my family' (CP: 29)

- the **vocative** is the form used to address someone. It is found with nouns but not with pronouns. An example is given in (15).

(15) *bhurg-yã=no,* *tumkã* *hi* *suvat* *kheɭ-p-a=c-ɛ̃* *mɵidan* *mhɵɳun*
child-OBL.PL=VOC.PL 2PL.OBJ this.F.SG place.F play-NMLZR-OBL=GEN-N.SG field.N QUOT

jay *kay* *"cilɖrens park"* *mhɵɳun* *jay?*
be.needed or Children's Park QUOT be.needed
'Children! Do you need this place as a playing field or as a "Children's Park"?' (Murkuɳɖe 2015: 5)

Having provided a basic overview of the functions of the cases given in Table 6, in Section 4 we turn our discussion towards the distribution of the two forms of the general genitive given in that table. We will argue that while the description of these markers in previous studies of Konkani is essentially correct, this description does not fully describe the present distribution of these two markers in the modern language. In Section 5 we then discuss the various ablative markers and will argue that one of these, *=cyan*, overlaps with both the ergative / instrumental and a somewhat enigmatic form which we refer to as the "pseudo-ablative / instrumental".

This discussion of the genitive and ablative forms will then play a central role in our discussion in Section 6 of two further cases, the selective and the elative, which are extremely common in all registers of Konkani, but which to our knowledge have so far escaped the attention of scholars of Konkani grammar.

4. The general genitive

In this section we discuss the status of what is referred to as the general genitive in Table 6 above, with the two enclitic markers =*c* or =*l*, followed by markers of agreement for gender, number and stem status of the head noun (= possessum). While the distribution of the kinship genitive is clear and only found with the meaning "belonging to [someone]'s household', as discussed in the previous section, the distribution of the general genitive markers =*c* and =*l* is less clear. The remainder of this section therefore deals with these two forms in some detail.

With respect to the distribution of the the "*c*- and *l*-genitives", as we refer to them here, Almeida (2004: 66) writes that the *c*-genitive can be used with all types of possessor NPs, whereas the *l*-genitive is restricted to nouns denoting personal names of human possessors. In a similar vein, Miranda (2003: 742) writes that the forms beginning with /l/ are found "when the preceding pronoun or noun denotes a human". In this section we look at the distribution of these two forms of the genitive as reflected in our corpus.

This corpus presently contains 21,628 words. It has two main subsets. The first part has been taken from a learner's manual of Konkani (Almeida 2004) and shows how Konkani is taught in Goa itself to non-Konkani speakers, so that the morphosyntactic structures that it contains can be considered to be what most native speakers of Goan Konkani view as prescriptively correct Standard Goan Konkani.[18] This part of the corpus consists of all but a few of the most basic introductory lessons found in Almeida (2004), including all stories and dialogues as well as example sentences in the exercises.[19] The second part of the corpus contains texts from a variety of sources. These include two spoken narratives, several short stories from children's books, and various sections of an academic text. (16)-(17) provide examples of the *l*-genitive from this corpus, where the possessor is the name of a person. Table 9 presents the statistics for the *c*- and *l*-genitives and the percentage of human possessors.

[18] However, this does not hold for orthography, which deviates somewhat in Almeida (2004) from standard practice. E.g., Almeida (2004) prioritizes morphological consistency in some cases where standard orthography follows the phonetic principle. For example, where the imperfective marker /ta/ follows a stem-final retroflex, e.g. *soḍ* 'leave', Almeida consistently writes this as <soḍta> / <सोडता>, preferring a consistent form of the imperfective marker, whereas Standard Goan Konkani orthography shows the assimilation of the initial /t/ of the imperfective to /ʈ/, i.e., <soḍʈa> / <सोडटा>, reflecting pronunciation.

[19] Only those exercises were excluded from our corpus where example sentences were not given in a final, "correct" form, as e.g. when the learner was expected to re-order the words of the sentence, or similar exercises.

Table 9: Human possessors and the genitive markers in our corpus

Learner's corpus	
Total number of genitive forms	
c-genitive	900 (98%)
l-genitive	15 (2%)
Total	**915**
Human personal names marked by the genitive in the learner's corpus	Total number: 41
Names followed by the *c*-genitive	26 (63%)
Names followed by the *l*-genitive	13 (32%)
Other human nouns followed by l-genitive	2 (5%)
Mixed "natural" corpus	
Total number of genitive forms	
c-genitive	419 (100%)
l-genitive	0 (0%)
Total	**419**
Human personal names marked by the genitive in the mixed corpus	Total number: 22
Followed by the *c*-genitive	22 (100%)[20]
Followed by the *l*-genitive	0 (0%)

(16) *uma, tumka tereja=l-i kombi avɖ-ũ na?*
Uma 2HON.OBJ Teresa=GEN-F.SG hen.F.SG be.pleasing-INF NEG.COP.PRS.3SG
'Uma, did you not like Teresa's hen (lit.: has Teresa's hen not pleased you)?'[21] (Almeida 2004: 28)

(17) *ai, nita=l-e ai=n suŋaʈ-ã=c-ĩ bhaj-ĩ ke-ll-ĩ.*
mother Nita=GEN-OBL mother=ERG prawn-OBL.PL=GEN-N.PL type.of.food-N.PL do-PST.PERF-N.PL
'Mother, Nita's mother made (lit.: had made) prawn *bhajas*.' (Almeida 2004: 31)

Two further lexical nouns referring to humans are also followed by the *l*-genitive marker, shown in (18) and (19). The possessum in (19) is omitted in the original text, as it is clear from context.

(18) *mhɵ=gel-ya iʃʈ-a=l-ɔ*[22] *aj jɵlmadis.*
1SG.OBL=KIN.GEN-OBL friend-OBL=GEN-M.SG today birthday.M.SG
'Today is the birthday of a friend of my family.' (Almeida 2004: 64)

[20] In the children's stories, an older man who plays a prominent role in helping the children is often referred to as kaka 'paternal uncle', which in this usage functions similarly to a personal name and is therefore counted as such here.

[21] The negative perfect is marked in (16) through the infinitive plus the present-tense negative copula, hence the lack of any explicit perfect morpheme in (16).

[22] *iʃʈ* 'friend' is also found once followed by the *c*-genitive in the learner's manual.

(19) *lok-ɑ̃=l-ɛ̃* *pɵɭe-vn* *hɑ̃vɛ̃=y* *dɵvɵr-l-ɔ.*
people-PL.OBL=GEN-N.SG look-CVB 1SG.ERG=ADD place-PST-M.SG
'Having watched the people's [behavior] I too placed [the present (M.SG) there].' ([Almeida 2004: 154)

If we expand "human personal names" to include "humans" in general, following Miranda (2003: 742), as in examples (18)-(19), this means that 15 out of 41, or 37%, of all human possessors take the *l*-genitive, while 63% of all cases with a human possessor are marked by the *c*-genitive. This confirms both Almeida (2004) and Miranda (2003)'s analysis of the *c*-genitive as unmarked with respect to the status of the possessor, whereas the *l*-genitive is restricted to human possessors. It is however also clear that even with nouns with human reference, the *c*-genitive is preferred, even in the learners' manual. As the *l*-genitive is found in our corpus exclusively in Almeida's (2004) learners' manual and never in the "natural" corpus, the use of the *l*-genitive is clearly optional in Standard Goan Konkani. Also, as these two genitive markers are interchangeable with human possessors, we consider them allomorphs of the same case, so that we assume one general genitive case.

Traces of the *l*-genitive are also found with a small number of adjectives which derive from postpositions or relational nouns, e.g., *vɵyl-* 'upper', which derives from *vɵyɵr* 'above' plus the *l*-genitive (< **vɵyɵr-l-* 'upper-GEN'), *bhayl-* 'outer' (< *bhayɵr* 'outside'), *faɻɵl-* 'posterior' (< *faɻ* 'back'), and *bhitɵrl-* 'inherent, intrinsic' (< *bhitɵr* 'inside'). This use of the *l*-genitive is no longer productive, and all of the above-mentioned forms are accordingly listed as separate lexical entries in Borkar et al. (2017), the standard dictionary for Goan Konkani. Similar comments hold for a few further adjectival forms which historically derive from the simple form of an adjective to which the *l*-genitive is added, with no notable semantic contribution. A few examples are given in (20).[23]

(20) *ɵʃɛ̃* / *ɵslɛ̃* 'like this; in this way'
tɵʃɛ̃ / *tɵslɛ̃* 'like that; in that way'
kɵʃɛ̃ / *kɵslɛ̃* 'how; in what way'.

Again, all of these forms are also listed as separate lexical entries in Borkar et al. (2017), underlining their status as lexicalized forms.[24]

In contrast, the *c*-genitive is entirely productive and, in addition to its use with all types of NPs to mark the possessor in complex NPs, it is also productively used to derive attributive forms from temporal and local deictic adverbials, such as *haŋ* 'here' (cf. e.g. example (13) above), *atɑ̃* 'now', *kal* 'yesterday', *tenna* 'then', etc.

[23] Note that /s/ before /e/, /ɛ/ or /i/ is regularly realized as [ʃ] and as [s] elsewhere in Konkani, so that the <s>~<ʃ> alternation in orthography in these examples is totally predictable and regular.

[24] This use of the *l*-genitive incidentally resembles that found in many Indo-Aryan and Munda languages of Jharkhand, in which the genitive marker came to be used as a focus marker, and later became so common in this function with certain words that it eventually lost its focusing function and essentially became part of the respective lexeme. Cf. e.g. Peterson (2010: 81-82; 2017: 561-568, especially the discussion at the bottom of p. 565).

In conclusion, we assume for Standard Goan Konkani a general genitive case which includes both the *c*- and *l*-genitives. With a possessor NP with human reference, the genitive may optionally be marked by *=l*, with no difference in meaning. The *l*-genitive is also found with a few other forms, where it has however now become part of the stem. Importantly, however, the *c*-genitives are the default forms and can always be used productively in the modern language.

5. The general ablative markers

Table 6 contains five different forms which are often mentioned as ablative markers in previous works. To these we can add the forms *=thavn* and *=sakun*. These are summarized in (21).

(21) *=san / =sun / =sakun / =savn /=than / =thavn / =cyan*

With the exception of *=cyan*, these ablative markers appear to be dialectal variants belonging to two main groups, i.e., *=san / =sun / =savn / =sakun* and *=thavn / =than*, respectively. They can all be viewed as regional / idiolectal variants of a single ablative case. These forms only mark ablative semantics. *=cyan* is different: To begin with, unlike the others, it transparently derives from the *c*-genitive with neutral, singular, oblique marking, i.e., *=c-ya*, plus the ergative / instrumental marker *=n*. On the other hand, the *s*-series of ablative markers in (21) seems to derive from the postposition *pasun*, while the *th*-forms are of a different source which is still unclear.

=cyan also has a different status than these other ablative markers, one which has previously been overlooked by many researchers (including by the present authors), namely that it can mark both an ablative NP but also an instrument. An example of an (unsuccessful) instrument function is given in (22) from Dalgado (2022: 115), who was perhaps the first to note this instrumental function of *=cyan*.

(22) *ta=cyan* *nəjɔ* *ja-l-ɛ̃.*
3SG.M.OBL=INST NEG.be.able COP-PST-N.SG
'It was not possible for him to do, he could not do (lit.: '[it] did not happen through/by him').' (Dalgado 2022: 115)

These (negative) abilitative constructions are typically expressed elsewhere in Indo-Aryan by marking the would-be agent by an ablative postposition, e.g., *se* 'from' in Hindi, which is then also an ablative / instrumental marker, and in many other Indo-Aryan languages the verb appears in the passive.[25] That *=cyan* in Konkani also marks ablative semantics is shown in (23).

[25] Cf. Hindi *mujh se* *nahīṃ* *so-yā* *ga-yā*
1SG ABL NEG sleep-PST.M.SG PASS-PST.M.SG
'I couldn't sleep, couldn't get to sleep (it was not slept by me).' (McGregor 1995[3]: 130)

(23) *jun-a=c-e* *pɵndra* *tark-e=cyan.*
June-OBL=GEN-OBL.F.SG fifteen date.F-OBL=ABL
'From the 15th of June [onwards].' (Almeida 2004: 51)

For this reason, we will refer to the marker *=cyan* henceforth as the ablative / instrumental marker, as it differs in this functional overlap from both the ergative / instrumental marker *=n* as well as the other ablative markers given in (21). For the sake of intelligibility, we will gloss it with the term that best fits the example it is found in.

=cyan may have been motivated through contact with Kannada, with which Konkani has been in contact for hundreds of years. Kannada has had an enormous impact on the morphology and syntax of Konkani, as until recently speakers of Konkani were generally lifelong bilinguals in both languages.[26] E.g., in Standard Kannada, there is an ablative / instrumental case marker, *=iṃda*, which follows a genitive-marked NP (cf. Zydenbos 2020: 97-98).

While the development of a postposition with instrumental and ablative functions and which takes an NP in the genitive is of course nothing out of the ordinary, the similarity between the two languages in this respect should not be entirely dismissed as coincidental, especially since there is similar evidence elsewhere for Kannada influence on the case system of Konkani: Cf. once again the form of the superessive *=r* and its unexpected allomorph *=cer*, with the genitive marker *=ce*, discussed in the previous sections. Similar to the instrumental / ablative, which takes the genitive, the locative case is marked in Kannada by adding the case marker *=alli* 'LOC' (< *alli* 'there') to the genitive-marked noun (cf. Zydenbos 2020: 98-100). As these are the only two case markers in both Standard Kannada as well as in Standard Goan Konkani which can or must take the genitive marker, Kannada influence on Konkani would seem to be a likely explanation of these two Konkani forms. This topic requires further study from a diachronic perspective.

Summarizing: In our final analysis of the Konkani case system, we give *=cyan* a distinct status as the ablative / instrumental case marker, and no longer consider it an allomorph of the general ablative, with which it only partially overlaps functionally. Instead, it can mark the ablative, unlike the ergative / instrumental marker *=n*, but it can also mark the instrumental, unlike the other ablative markers in (21).

One last form must also be mentioned with respect to the ablative, namely the somewhat enigmatic morph *=lyan*. While not a productive ablative or instrumental marker, *=lyan* has the same underlying structure as *=cyan*, differing from the latter only in that it derives from the *l*-genitive instead of the *c*-genitive, i.e., *=l-ya* + the ergative / instrumental marker *=n*.[27] However, unlike *=cyan*, which can mark ablative and instrumental semantics, *=lyan* is not a productive marker from a synchronic perspective.

It is difficult to ascertain from the modern language whether *=lyan* was ever productively used, and further research on this form in older Konkani is necessary. *=lyan* is only

[26] See in this respect e.g. Nadkarni (1975), Peterson (2022) and Peterson & Chevallier (2022).

[27] In our view, that fact that these two markers have an identical underlying form, one deriving from the *c*-genitive and one from the *l*-genitive, further suggests that *=cyan* has been structurally copied from Kannada with purely Konkani morphology, and suggests an earlier stage where both forms competed with one another. Again, this topic requires further study.

found with three lexemes in our corpus, all of which have lexicalized with somewhat unpredictable meanings: *kəḍ* 'side', *faṭ* 'back' and *vəyər* 'above'. When combined with *=lyan* these yield *kəḍlyan* 'from', *faṭlyan* 'behind' and *vəylyan* 'on top of that / in addition to that' but also 'from above' (cf. Borkar et al., 2017: 570). In addition to 'from', *kəḍlyan* can also have an instrumental meaning, as (24) shows.

(24) *bhayər mej-a=r bəs-la* ***tya*** ***mənf-a*** ***kəḍlyan*** *bhər-un ghe.*
outside table-OBL=SUPESS sit.down-PRS.PERF that.OBL man-OBL INSTR fill.out-CVB AUTOBEN.IMP
'Have it filled out **by the man** seated at the table outside.' (Almeida 2004: 113)

Of the above-mentioned forms, only *kəḍlyan* 'from / through (INST)' and *vəylyan* 'from above' now show a direct connection to the ablative / instrumental, suggesting that if *=lyan* was once productive, it has since been entirely replaced in this function by the form *=cyan*.

There is one environment, however, in which *=lyan* and the *l*-genitive are both still commonly found with an ablative / instrumental or a genitive function, respectively, irrespective of animacy. These are discussed in the following section. To facilitate the discussion there, in the following we refer to the marker *=lyan* as the "pseudo-ablative / instrumental" form, to differentiate it from the "real" ablative / instrumental marker *=cyan*. As the *l*-genitive is still in use, we will simply refer to this form in the following as the "*l*-genitive".

6. The selective[28] and elative cases

Our corpus contains numerous examples of two further case markers which are very common in both data sets of our corpus but which to our knowledge have not received mention in any grammar or learner's manual of Konkani until now. Both derive historically from case stacking. The first of these, the selective, derives from the inessive singular marker *=nt* followed by the *l*-genitive marker. It refers to an entity which is located within a larger entity, from which it is "selected". The second, the elative, derives from the inessive singular marker *=nt* and the pseudo-ablative / instrumental allomorph *=lyan*. It refers to movement out of an entity or group of entities.

Examples (25)-(32) illustrate the forms and functions of these two cases. The selective case is restricted to attributive function ((25)-(28)). The elative is tendentially found in attributive function (e.g., (29)-(30)), although it can also function adverbially at the sentence level, as (31)-(32) show.

Selective: *=ntl-*

(25) *atã* ***don*** ***hǝjar*** ***don-a=ntl-ɔ*** ***ek*** ***maḍ*** *lag-lǝl-ɔ*
now two thousand two-OBL=SEL-M.SG one coconut.tree.M bear.fruit-PST.PERF-M.SG

[28] The term *selective* has been inspired by the use of this term in Wagner-Nagy (2019: 521), although the status and function of our use of the term differs considerably from that in Wagner-Nagy's grammar. We thank Chris Lasse Däbritz for calling this term to our attention.

na.
NEG.PRS.3SG
'Now not **one coconut tree among those planted in 2002** bore fruit.' (CP: 18)

(26) ***tɛ*** ***prøt-i=ntl-ĩ*** *pan-ã* *bør-ĩ=c=ʃ-ĩ* *vac-ũ* *ye-tal-ĩ*
that.OBL copy-OBL=SEL-N.PL page-N.PL good-N.PL=FOC=APPROX-N.PL read-INF be.able-IPFV.PST-N.PL

øʃ-ɛ̃ *mhaka* *dis-l-ɛ̃.*
such-N.SG 1SG.OBJ be.seen-PST-N.SG
'**The pages in that copy** were easily readable, so it seemed to me.' (Miranda 2019: ix)

(27) ***gãv-a=ntl-ɛ*** ***lok*** *kasl-ɛ* *dhønd-ɛ* *kør-tat?*
village-OBL=SEL-M.PL people of.what.type?-M.PL profession-M.PL do-PRS.PL
'What professions do **the people in the village** do?' (Almeida 2004: 138)

(28) ***hya*** ***viʃøy-a=ntl-ɛ*** ***coɖ*** ***mahiti*** ***khatir*** *mhaka* *pi.ke.rajʃekhør*
this.M.OBL subject-OBL=SEL-F.SG.OBL much information.OBL for 1SG.OBJ P.K.Rajshekar

hãɳi *sɵpadit* *ke-ll-ɛ̃* *kønnøɖ* *jønpød* *møhabharøt* *jay* *as-l-ɛ̃.*
3HON.ERG edited do-PST.PTCP-N.SG Kannada people Mahabharata be.wanted COP-PST-N.SG
'**For more information on this subject**, I wanted the Kannada People's Mahabharata edited by K.P. Rajshekar.' (Miranda 2019: xiii)

Elative: *=ntlyan*

(29) ***payp-a=ntølyan*** ***udøk*** *soɖ-un* *ekømek-ã=k* *bhijovn bhij-øy-un* *møj-e=n*
pipe-OBL=ELA water release-CVB RECP-OBL.PL=OBJ get.wet-CAUS-CVB fun-OBL=INST.SG

dhuɭvøɖ *mønøy-l-i.*
sprinkling.of.colors.during.Holi.F celebrate-PST-F.SG
'[The boys and girls] turned on (lit. 'released') the faucet ('the **water from in the pipe**'), got each other wet and joyously (= with fun) celebrated the throwing of colors of Holi.' (Murkuɳḍe 2015: 10)

(30) *tãɳi* *mhøje* *khatir* ***bhã.pra.sõ.sõ=c-ya*** ***grønthaløy-a=ntølyan*** ***tya***
3.HON.ERG 1SG.POSS.OBL for B.O.R.I.=GEN-OBL.N library-OBL=ELA that.OBL

grønth-a=c-i ***foʈokɔpi*** *kør-p-a=c-i* *vevøstha* *ke-l-i.*
book-OBL=GEN-F.SG photocopy.F do-NMLZR-OBL=GEN-F.SG arrangement.F.SG do-PST-F.SG
'She arranged to make **a photocopy of that book from the B.O.R.I.'s library** for me.' (Miranda 2019: xii)

(31) *pracin* *koŋkɳi* *bharøt-a=c-i* *khøbør* *mhaka* *pøylĩ* *ɖa.* ***jujhe***
ancient Konkani Mahabharat-OBL=GEN-F.SG news.F.SG 1SG.OBJ first Dr. Jorge

pirer-ã=c-ya ***børp-a=ntølyan*** *me[-[-i.*
Pereira-HON.OBL.HON=GEN-OBL writing-OBL=ELA meet-PST-F.SG
'I first came across news of the old Konkani Bharat **from [a mention] in a writing of Dr. Jorge Pereira**.' (Miranda 2019: ix)

(32) *ta=c-ya* ***burak-a=ntlyan*** *koɳ=ɵy* *barik* *sarik* *jiv* *bhitər* *sər-ũk*
that.OBL=GEN-OBL hole-OBL=ELA who?=ADD small ECHO creature inside move-INF

ʃək-tal-ɔ.
be.able-PST.IPFV-M.SG
'Any tiny creature could come in through (= **from inside**) **that hole**.' (Nayək 2017: 6)

In combination with the demonstratives *hɛ̃* 'this' and *tɛ̃* 'that', the respective base forms for these two cases are realized as either *hatunt-*/*tatunt-*, or *hantũ-*/*tantũ-*, respectively, to which the *l*-genitive or *=lyan* is added; cf. example (33).

(33) *ami* *jenna* *kaka* *vaŋɖa* *hi* *suvat* *nivəl* *kər-tal-ɛ,* *tenna*
1PL when.CRL uncle together.with this.F.SG place.F clean do-PST.IPFV-M.PL then

yẽvk
hatuntl-ɔ ***koɳ=uc*** *adar-a=k* *ye-ũk* *naʃ-ill-ɔ.*
this.SEL -M.SG who=FOC help-OBL=OBJ come-INF NEG.COP-PST.PERF-M.SG
'When we were cleaning this place together with Uncle, then **no one from among them** came to help (lit.: **who among these** had not come to help).'[29] (Murkuɳɖe 2015: 11)

The most likely reason that none of the grammatical works that we consulted deal with these forms[30] is that these appear at first glance to be typical examples of case stacking, in which a genitive or (pseudo-)ablative / instrumental case marker is added to an inessive-marked NP with entirely predictable semantics. For example, the complex NP *paypantəlyan udək* in (29) is semantically transparent: water (*udək*) from inside the pipe (*paypantəlyan*); similar comments apply to all other examples in our corpus, as examples (25)-(33) above show. The main question is thus whether these forms are still analyzable as case stacking in the modern language or whether they have fossilized and become case markers in their own right. In the following we show that the facts clearly indicate that both of these markers have indeed fossilized and must be viewed as simple case markers from a synchronic perspective.

Assuming for the moment that the elative and the selective are in fact case stacking in the modern language, we would expect three things, namely:

1. that both the *l*- and *c*-genitive markers are found here with roughly the same distribution as described above in Section 4;
2. that the ablative / instrumental marker *=cyan* should be possible with the elative, since of the two only *=cyan* is productively used in the modern language whereas *=lyan* is not, and:
3. that the inessive marker has the form *=nt* with a singular noun and *=ni* with a plural noun (cf. again Table 6).

[29] Note that *koɳ=uc* [who=FOC] plus a negated verb means 'no one'.

[30] The selective is found once in Dhongde (2022: 157, example (99)) in the form *goy-a-ntl-i*, which Dhongde glosses as 'Goa-OBL-PP:in.F.S', unfortunately without further comment or translation of the form.

However, as we now show, none of these three conditions hold.

With respect to the first point, note that only the *l*-based forms of the genitive are possible in the selective, regardless of the status of the possessor as [±human], so that Point 1 does not hold (cf. e.g. examples (25)-(28) and (33)). This also holds for the elative, where only *=lyan* can occur; thus, Point 2 also does not hold. That is, the forms **=ntcɛ̃* and **=ntcyan* are not grammatical in Standard Goan Konkani,[31] and these two units do not show the distribution we would expect for either the genitive or the ablative / instrumental.

With respect to Point 3: Only the "singular" form *=nt* is found in the selective and the elative, even with a plural noun. Consider the lexeme *bhas* 'language.F' in examples (34) and (35). In the singular, *bhas* has the oblique stem *bhaʃe* while in the plural it has the oblique stem *bhasɑ̃*, written *bhasa* before /nt/. In both (34) and (35), these two stems, both singular and plural, are followed by the same form of the selective, namely *=ntəl-*.

(34) ***dər*** ***ek-e*** ***bhaʃ-e=ntəl-ya*** ***pracin*** ***sahity-a=k*** *ayc-ya*
every one-OBL.F.SG language.F-OBL.F.SG=SEL-OBL ancient literature-OBL=OBJ of.today-OBL

kaḷ-a=nt *ek* *agḷ-ɛ̃=c* *mhətvə* *prapt* *ja-ta.*
time-OBL=INESS one special-N.SG=FOC significance acquired become-PRS.3SG
'The ancient literature of every language obtains a special significance in today's time (lit.: **to the ancient literature which is in every single language** … a special significance is acquired).' (Miranda 2019: viii)

(35) *pracin* *məraṭhi,* *sɔ̃skṛt ,* *kənnəḍ* *ani* ***her*** ***bhas-a=ntəl-yo***
ancient Marathi Sanskrit Kannada and other language-OBL.PL=SEL-F.PL

məhabharət-a=c-yo *prət-i…*
Mahabharata-OBL=GEN-F.PL copy-F.PL
'… copies of the Mahabharata in Old Marathi, Sanskrit, and Kannada and **in those in other languages** …' (Miranda 2019: viii)

In (34)-(35) it is only the oblique stems (*bhaʃe* (SG) and *bhasɑ̃* (PL)) that indicate number, not the case markers. If the selective were indeed case stacking in the modern language, we would expect the **ungrammatical** plural form **bhas-a=ni=l-yo* in (35), with the plural inessive marker *=ni*, instead of the grammatical form *bhas-a=ntəl-yo*, as *=nt* as a productive case marker signals the **singular** inessive. As the expected plural marker *=ni* of the inessive cannot appear here together with the plural stem but only the "singular" form *=nt*, the form *=ntəl-* cannot be viewed as case stacking in the modern language. Rather, the two historical case markers INESSIVE + GENITIVE have fused to form a new case. Similar comments hold for the elative.

In fact, it is only in those declensions where the singular and plural oblique stems differ with respect to their final vowels, such as *bhaʃe* (SG) vs. *bhasɑ̃* (PL) in (34)-(35),

[31] These forms do, however occur in some non-standard varieties of Konkani, although not in the Goan Standard. This topic requires further study, as this could mean that there is still productive case stacking in those non-standard varieties. We will deal with this topic in a later study, as detailed data on these dialects are not yet available. However, as the discussion above shows, in the Goan Standard variety we must assume that the elative and the selective are cases in their own right from a synchronic perspective.

that the singular and plural of these cases can be distinguished morphologically. For example, the majority of masculine and neuter stems end in *-a* in the oblique singular and *-ã* in the oblique plural (which is written as <a> before <nt>), e.g., *put* 'son.M.DIR.SG' vs. *puta* 'son.M.OBL.SG' / *putã* 'son.M.OBL.PL' and *ghər* 'house.N.DIR.SG' vs. *ghəra* 'house.N.OBL.SG' / *ghərã* 'house.N.OBL.PL'. These nouns therefore cannot formally express singular vs. plural in the elative and the selective.

These facts, taken together, show that the selective and the elative must be viewed as fully grammaticalized case markers in Standard Goan Konkani and are probably of recent origin.

7. Analysis: How many cases are there in Standard Goan Konkani – and why?

Table 10 summarizes the results of our analysis with respect to the case forms of Standard Goan Konkani. We assume the 13 distinct cases listed in that table. In this section we justify our decision for assuming each of these cases. We recognize that the status of these markers as distinct cases or as allomorphs of more general cases is to some extent dependent on the theoretical persuasion and/or practical concerns of individual researchers; we therefore end this section by showing how our own analysis compares with both a minimal and a maximal analysis of this system.

With respect to its status as a case in its own right, the only candidate in Table 10 which is truely uncontroversial is the **vocative**; it has a unique marking pattern, with zero marking of the oblique stem in the singular and marked by *=no* in the plural, and a clearly defined function, namely addressing someone. While the vocative does not fit the usual definition of a case, as was discussed above in Section 2, we follow the argumentation in Blake (1994: 9) and view the vocative as a case in Konkani, due to its structural similarities with other cases.

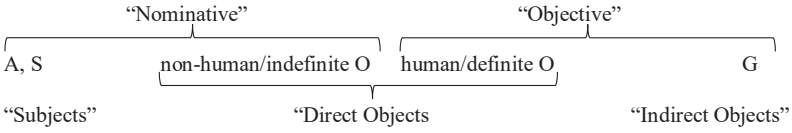
Table 10: The revised case system of Konkani

Case	Singular	Plural
Nominative (= direct stem)	*Ø*	
The following enclitic case markers attach to the oblique stem:		
Objective	*=k*	
Ergative / Instrumental	*=n*	*=ni*
Ablative / Instrumental	*=cyan*	
Locative cases		
Inessive	*=nt (=n)*	*=ni*
Superessive	*=r / =cer*	
Kinship essive	*=ger*	

Ablative cases		
Ablative (general)	*=san / =sun / =sakun / =savn /=than / =thavn*	
Elative	*=ntlyan*	
Genitive cases		
Genitive (general) – *l*-forms are restricted to human possessors, *c*-forms are unmarked in this respect	*=cɔ / =ci / =cɛ̃* or *=lɔ / =li / =lɛ̃*	
Kinship genitive	*=gɛlɔ / =geli / =gɛlɛ̃*	
Selective	*=ntlɔ / =ntli / =ntlɛ̃*	
Vocative	Ø	*=no*

The status of the **nominative** and the **objective** as distinct cases requires some comment. Although A and S (i.e., "subjects") in the present tense all appear in the nominative, and definite, human direct objects (O) are marked as objective, non-human and indefinite object NPs are zero-marked, i.e., they have the same form as the nominative. Furthermore, indirect objects (G) receive the same formal marking as definite and human O's, i.e. the objective case. Diagram 1 illustrates this distribution.

Diagram 1: The distribution of the formal markers of "subjects", "direct objects" and "indirect objects" in Konkani



For this reason, many researchers of South Asian languages with similar differential object marking patterns differentiate between an accusative and a dative case so that not only subjects and indirect objects, but also direct objects are assigned to a single case, at least in name. Alternatively, our objective case is divided into two distinct cases, the "accusative" and the "dative" which have the same marker.[32]

We follow the formal marking and assume two cases here, the nominative and the objective, as shown in Diagram 1. As this pattern is quite common crosslinguistically, with shared marking for S, A and non-human/indefinite O as opposed to G and human/definite O, we prefer to assume the presence of only two cases here, not three, although other researchers may analyze this differently.

The status of the **ergative / instrumental** as a unique case would seem straightforward, although it should also be mentioned that the same marker does cover two different semantic roles, namely an agent and an instrument. As such, some might chose to consider them two different cases, among other reasons as they can co-occur in the same sentence. However, we again prioritize the formal marking and consider this to be one

[32] Cf. e.g. the discussion in Butt (1995: 17-19) on Hindi and the literature cited there.

case.[33] Similar arguments hold for the **ablative / instrumental**, which we similarly recognize as a distinct case due to its form, despite its overlap with the ergative / instrumental on the one hand and with the general ablative markers on the other, since its specific semantics differ from that of both the ergative / instrumental and the purely ablative markers.

The three locative cases in Table 10 also seems straightforward, in that we have three different markers, one denoting location in a place (**inessive**), on something (**superessive**), or at the home of someone's family (**kinship essive**). We therefore consider these to be three different locative cases. Note however that Dhongde (2022: 55) groups all of these together as one overarching locative case, so that our analysis is not entirely uncontroversial.

The **general ablative** in our analysis is clearly a separate case, one which can be marked quite differently according to the region and personal preference of the speaker, although these are all allomorphs in free variation and all mark the same ablative function. The **elative** in our view is also a separate case; while its semantic relation to the general ablative is undeniable, it is at the same time distinct from this in that it refers to movement away from a temporal or locative source within a specific set of referents. The latter part of this definition is clearly lacking from the general ablative, which merely refers to movement away from a particular entity or time.

The genitive cases also present various problems with respect to the total number of cases of Goan Konkani. We assume that the *c*-genitive and the *l*-genitive together form one **general genitive** case which incorporates one NP or other unit (e.g., postpositions or adverbs) in attributive function into a larger NP: The difference between the two is the status of the possessor NP, either [+human] for the *l*-genitive or unspecified for this feature with the *c*-genitive, however, as the *l*-genitive is optional and can always be replaced by the *c*-genitive, we view these as subtypes of a more general genitive case.

The relationship of the **kinship genitive** and the **selective** with the **general genitive** is somewhat less straightforward. The main difference between these three markers, which are all used exclusively in attributive function, is the status of the possessor – i.e., kinship, a specific group of individuals of which the head noun is one, and all other nouns, respectively. Again, we give precedence here to form and view these three markers as three different cases. The etymologies of these forms are also relevant here: The selective derives from case stacking involving two different case markers with two different semantic roles, namely the inessive and possessor (in its broadest sense), both of which are still present in the semantics of the selective, although the marker is now invariable and can no longer be viewed as case stacking. In our view, the unique semantics and form suffice to qualify this category as an independent case.

[33] Recall however from Section 3.2 above that in the spoken language, the ergative / instrumental singular and the inessive singular are both commonly realized as *=n*, and both cases are regularly marked by *=ni* in the plural, so that in the speech of these speakers, the ergative / instrumental and the inessive are not distinguished in either the singular or the plural, which thus complicates this picture somewhat. Again, as we are dealing here with the standard dialect, we do not pursure this topic further here.

Similar arguments hold for the analysis of the kinship genitive: It seems to us that the semantic distinction between a general "possessor" and the members of a household as a possessor is a cultural distinction which must be recognized as a distinct category. This category likely derives from *ghər* 'house' plus the *l*-genitive marker, i.e., **ghər=l-* 'house=GEN-", where the /r/ of /ghər/ is lost before /l/, although this awaits confirmation.[34] From a cultural perspective, the fact that a distinction between **ghər=l-* and *=l-* was salient enough not only to be made, but also that **ghər=l-* was used so regularly with reference to a household that it evolved further phonologically to the point where its original form can only be inferred, strongly suggests that this difference is culturally important enough to consider it a separate case. This analysis is further strengthened by the presence of a kinship essive, discussed above, which for independent reasons we also classifiy as a separate case. We recognize, however, that opinions on this issue may differ among researchers.

Table 11 compares our results with a minimal and a maximal analysis of the Konkani case system. Our analysis of 13 cases in total – a rather high number when compared with the case systems of most other New Indo-Aryan languages – is presented in the center of the table. To the left, the minimal interpretation is given, with seven cases and the maximal interpretation with 16 cases, given on the right. There are of course other possible intermediate interpretations, however we believe these three interpretations suffice to illustrate the major underlying difficulty, namely that even after having identified the relevant forms of the language and their primary functions, there still remains the issue of deciding if the respective form is a case marker (as opposed to a postposition) and also when the meanings of two or more categories differ enough to consider them to be different cases, and when they are still close enough to be considered allomorphs of a more general case.

8. Summary and outlook

In this study we analyze the case system of Standard Goan Konkani, an Indo-Aryan language spoken throughout a narrow strip of land on the west coast of India. After defining *case* for the purpose of this study and how we structurally distinguish enclitic case markers from postpositions, we give a brief overview of the case systems proposed by various researchers over the past 100 years, ranging in size from six to eleven cases. We note however that it is not always entirely clear from such descriptions what counts as a separate case for the individual authors, as different forms are sometimes given under one heading without specifying their exact function or status.

[34] Cf. e.g. once again **vəyər=l-* 'top=GEN-' > *vəyl-* 'upper' in Section 4. Similarly, the kinship essive seems to derive from **ghər* 'house' or the superessive form **ghər-er*. Further work is necessary.

Table 11: The case system of Konkani – three possibilities

Minimal system	Our analysis	Maximal system
Nominative	Nominative	Nominative
Objective	Objective	Accusative
		Dative
Ergative / Instrumental	Ergative / Instrumental	Ergative
		Instrumental
	Ablative / Instrumental	Ablative / Instrumental
Locative	Inessive	Inessive
	Superessive	Superessive
	Kinship essive	Kinship essive
Ablative	General ablative	General ablative
	Elative	Elative
Genitive	Genitive	General genitive
		Human genitive
	Kinship genitive	Kinship genitive
	Selective	Selective
Vocative	Vocative	Vocative
7 cases	**13 cases**	**16 cases**

What this overview of past interpretations of the Konkani case system also shows is that choosing the most appropriate descriptive level for discussing case in Konkani is anything but trivial, and the answer to the question of how many cases Konkani "really has" will differ from researcher to researcher, depending on the respective researcher's theoretical persuasion and practical considerations.

We also provide a detailed discussion of these different markers and their functional range and discuss two further cases which have gone virtually unnoticed until now in grammatical discussions of Konkani, despite their very common use. These are the selective and the elative, which derive from earlier case stacking but which must be considered case markers in their own right in the modern language. We come to the conclusion that Standard Goan Konkani can best be described as having 13 productive cases, quite a large number for a New Indo-Aryan language, and conclude our discussion by comparing our own analysis with a minimal and a maximal analysis of the Konkani case system, showing how the number of cases can plausibly be argued to range anywhere from seven to 16, depending on the respective researcher's theoretical and practical assumptions.

The discussion of the previous pages also shows that the case system of Standard Goan Konkani is in the midst of a highly dynamic process of reanalysis and change and has likely been so for some time, with traces of old case forms co-existing with trans-

parent newer forms which show partial functional overlap with other forms, but also new case markers with unique semantics. Given Konkani's intimate historical relationship with Kannada, its much larger Dravidian neighbor to the South and East which has immensely affected its morphological and syntactic development over the past centuries, a closer comparison of the case systems of these two languages will undoubtedly yield further insights into these dynamics, as will a closer comparison with Marathi, Konkani's larger and better known sister language to the North, from which it separated several centuries ago.

Conflict of Interest: The authors declare no conflict of interest
Authors' Contributions: The authors confirm sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Abbreviations

1, 2, 3	– persons	GEN	– genitive	PART	– participle
ABL	– ablative	HON	– honorific	PERF	– perfect
ADD	– additive focus	IMP	– imperative	PL	– plural
ADJVZR	– adjectivalizer	INESS	– inessive	PP	– postposition
AUTOBEN	– autobenefactive	INF	– infinitive	PST	– past
CAUS	– causative	INST	– instrumental	QUOT	– quote
COP	– copula	IPFV	– imperfective	RECP	– reciprocal
CRL	– correlative marker	KIN.ESS	– kinship essive	REFL	– reflexive
CVB	– sequential converb	KIN.GEN	– kinship genitive	SG	– singular
DAT	– dative	M	– masculine	SEL	– selective case
DIR	– direct stem	N	– neuter	SUPERESS	– superessive
ELA	– elative	NEG	– negative	V2	– "vector verb" expressing Aktionsart
ERG	– ergative	NMLZR	– nominalizer		
F	– feminine	NOM	– nominative	VOC	– vocative
FOC	– focus	OBJ	– objective case		
FUT	– future	OBL	– oblique stem		

References

Almeida, Matthew, S.J. 1989. *A description of Konkani.* Miramar, Panaji, Goa: Thomas Stephens Konknni Kendr.

Almeida, Matthew, S.J. 2004. *Koŋkṇī muḷāvɔ kors – Konkani basic course*. (Thomas Stevens Konknni Kendr, 16). Alto Porvorim, Goa: Thomas Stephens Konknni Kendr.

Almeida, Matthew, S.J. 2012. *Konknni course in Roman script: Romi lipient Konknni kors.* Second edition. [TSKK Publications – 20]. Alto Porvorim, Goa: Thomas Stephens Konknni Kendr.

Blake, Barry J. 1994. *Case.* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Borkar, Suresh J. & Thali, Mukesh P. & Ghanekar. Damodar K. 2017[9]. *Rajhauns new generation Konkani-English illustrated dictionary.* Panaji, Goa: Rajhauns Sankalpana.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu.* Stanford, CA: CSLI.

Dalgado, Sebastião Rodolfo. 2022. *Grammar of Konkani language: Original Portuguese medium (Gramática da língua concani).* Co-Translators-cum-Editors: de Ataíde, Mousinho & de Menezes, Cristo. Panjim, Goa: Broadway Publishing House.

Dhongde, Ramesh Vaman. 2022. *Konkani.* (London Oriental and African Language Library 2). Amsterdam–Philadelphia: Benjamins.

Eberhard, David M. & Simons, Gary F. & Fennig, Charles D. (eds.). 2023. *Ethnologue: Languages of the world.* Twenty-sixth edition. Dallas, Texas: SIL International. (http://www.ethnologue.com) (Accessed: 2023-07-14)

Ghatage, A.M. 1963. *Konkani of South Kanara.* (A Survey of Marathi Dialects 1). Bombay: State Board for Literature and Culture.

Ghatage, A.M. 1965. *Kuḍāli*. (A Survey of Marathi Dialects II). Bombay: State Board for Literature and Culture.

Ghatage, A.M. 1966. *Kuṇabī of Mahāḍ*. (A Survey of Marathi Dialects III). Bombay: State Board for Literature and Culture.

Ghatage, A.M. 1968. *Konkani of Kankon*. (A Survey of Marathi Dialects V). Bombay: State Board for Literature and Culture.

Ghatage, A.M. 1972. *Gāwḍi*. (A Survey of Marathi Dialects VIII). Bombay: State Board for Literature and Culture.

Karapurkar, Pushpa Pai. 1968. *Descriptive analysis of Gawdi.* Poona: Deccan College. (Doctoral dissertation.)

McGregor, R.S. 1995[3]. *Outline of Hindi grammar with exercises.* Third edition, revised and enlarged. (Oxford India Paperbacks). New Delhi: Oxford University Press.

Miranda, Rocky. 2003. Konkani. In: Cardona, George & Jain, Dhanesh (eds.), *The Indo-Aryan languages.* (Routledge Language Family Series), 729-765. London–New York: Routledge.

[Miranda, Rocky V.] Mirand, Rɔki Va. 2019. *Pracin koŋkṇi bharət. Khəṇḍ 1: vəḷəkh. koŋki ənuvad: damodər ghaṇekar, prəkaʃ thəḷi, sɛbi fernandiʃ, rɔki va. mirand.* [The Old Konkani Bhārata – Volume 1: Introduction: Konkani translation: Ghanekar, Damodar & Thali, Prakash & Fernandes, Sabi & Miranda, Rocky V.]. Məḍgāv, Gõy: Asmitay Prətiṣṭhan.

Murkuṇḍe, Dipa Dinkər. 2015. *Itsa thẽy marg*. [Where there's a will, there's a way.] Naveli-Salset-Gõy: ʃri Həri Prəkaʃən.

Nadkarni, Mangesh V. 1975. Bilingualism and syntactic change in Konkani. *Language* 51(3). 672-683.

Nayək, Gita Narayəṇ. 2017. *bappalyo kaŋyo.* Məḍgāv: əʃina prəkaʃən. [Nayək, Mrs. Gita Narayəṇ. *Father's stories.*

Peterson, John. 2010. Language contact in Jharkhand: Linguistic convergence between Munda and Indo-Aryan in eastern central India. *Himalayan Linguistics* 9(2). 56-86. https://doi.org/10.5070/H99223479

Peterson, John. 2017. Jharkhand as a 'linguistic area': Language contact between Indo-Aryan and Munda in eastern-central South Asia. In: Hickey, Raymond (ed.), *The Cambridge handbook of areal linguistics.* (Cambridge Handbooks in Language and Linguistics), 551-574. Cambridge: Cambridge University Press.

Peterson, John. 2022. A sociolinguistic-typological approach to the linguistic prehistory of South Asia: Two case studies. *Language Dynamics and Change* 12(2). 224-273. https://doi.org/10.1163/22105832-bja10018

Peterson, John & Chevallier, Lennart. 2022. Towards a typology of negation in South Asian languages. *Bhāṣā, Journal of South Asian Linguistics, Philology and Grammar Traditions* 1(1). 17-62. http://doi.org/10.30687/bhasha/8409-3769/2022/01/005

Peterson, John & Mopkar, Govind. 2021. Past habitual actions as relative future? On an unexpected use of the Konkani future participle and its likely origin. *Lingua Posnaniens* 63(2). 33-54. https://doi.org/10.14746/linpo.2021.63.2.2

Wagner-Nagy, Beáta. 2019. *A grammar of Nganasan.* (Grammars and language sketches of the world's languages. Indigenous languages of Russia). Leiden–Boston: Brill.

Zydenbos, Robert J. 2020. *A manual of Modern Kannada*. Heidelberg: CrossAsiaeBooks. https://doi.org/10.11588/xabooks.736

DOI: 10.14746/linpo.2024.66.1.4

REVIEWS

Patience Epps and Lev Michael (eds.). 2023. *Amazonian languages*; *Language isolates I: Aikanã to Kandozi-Chapra* (pp. i-lxi, 1-657 + "Index" [24 pp. not numbered]); *II: Kanoé to Yurakaré* (pp. i-xiii, 659-1352 + "Index" [26 pp. not numbered]); *An international handbook* (Handbooks of Linguistics and Communication Science // Handbücher zur Sprach- und Kommunikationswissenschaft (HSK) series, vols. 44.1-2 of HSK 44). Berlin–Boston: De Gruyter Mouton

Alfred F. Majewicz

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

The Amazon basin is the least known
and least understood linguistic region in the world.

The above motto is the very first (save xxviii pp. of the front matter) sentence of Dixon and Aikhenveld 1999 (:1), a book which on its very first page ([i]) states that "The Amazon basin is arguably both the least-known and the most complex linguistic region in the world today", reiterating it on the fourth (back) cover of the edition ("Paperback Re-issue") at this writer's hand. Discussing reasons for such a "state of the art", the editors point, among others, to the fact that "[...] the standard of scholarship in South American linguistics is not high" [...] but they "[...] must add that there are notable exceptions [...] – a number of descriptive studies that achieve a high standard of clarity and explanation" and refer in that context specifically to the *Handbook of Amazonian languages* (Derbyshire & Pullum) evaluating it to be "a most worthwhile enterprise [...] so far [...]" (*ib.*, p. 3).

The appearance of Derbyshire & Pullum's first volume of their *Handbook* in 1986 aroused immediate interest of the editors of *Lingua Posnaniensis* (LPos) evidently aware of the situation in South American linguistic research as described by Dixon and Aikhenveld and quoted above, and a review of the volume was commissioned just after its

release in print. Submitted to the editorial board for publication in 1987[1], due to an unfavorable coincidence of undesirable circumstances it could see the light of the day only over three years later (LPos 31 (1990), 120-2). Derbyshire & Pullum's *Handbook* was initially planned as a series of three volumes (1986: v) and a review of both volumes 2 (1990) and 3 (1991), also commissioned by LPos, was published in 1992 in a newly established journal (LOS 1, 236-8)[2].

Derbyshire & Pullum 1986 (xiii + 642pp.) includes, apart from "Introduction" and "Part II: Word order and typological studies", "Grammatical sketches" of four languages: Apalai (pp. 33-127), Canela-Krahô (128-99), Pirahã (200-325), and Urubu-Kaapor (326--403). In Derbyshire & Pullum 1990 (x + 474) one finds two such sketches (in fact, pretty extensive grammars) of Sanuma (15-248) and Yagua (249-474), and in Derbyshire & Pullum 1991 (xii + 517) – also two "sketches" of Macushi (23-160) and Paumarí (161-352) alongside with "A classification of Maipuran (Arawakan) languages" and a "Cumulative index to volumes 1-3"[3].

The names of the editors of the 1986ff. *Handbook...* – Derbyshire (1924-2007) and Pullum (born 1945) – appear in literature frequently in pair (side by side) and, surprisingly for many to learn, it was the latter, a linguist specializing in general linguistics and research focusing on the... English language, who actually made the former, an accountant-incidentally-turned-SIL-worker among Hixkaryána Indians in northern Brazil[4], an academic linguist specializing in Carib languages[5]. And it was the Hixkaryána language and its OVS, considered "nonexistent", word order that made both Derbyshire and Pullum famous in the linguistic world and greatly helped advance studies on "Amazonian languages" as exemplified by the above quoted Dixon & Aikhenveld 1999, 660-page volume on Jarawara (Dixon 2004 in its 2011 paperback edition), equally bulky MGL *Aguaruna*

[1] Cf. the notice of receipt of the typescript: <*Allatum die 2 mensis Januarii a. 1987*>, LPos 31 (1990), 122.

[2] Due to a drastic social, economic and political crisis in Poland, the country found itself on the verge of collapse and the publication of *Lingua Posnaniensis* (in existence since 1949; its vol. 1 included Loukotka's paper on "some unknown languages of... South America") was under threat of liquidation of the title. LOS was conceived and created to dispose of growing piles of academically valuable and attractive materials submitted and already accepted for publication in the endangered LPos and – to save faces. The situation happy-ended for Poland with its dramatic change of political system and NATO and EU membership as well as for LPos (survival and continuation) and LOS (17 vols. in two series published between 1991-2014). Loukotka's texts appeared also in LPos 2 (1950), 3 (1951), 5 (1955), and 6 (1957). He was renowned as the author of the 1968 "catalogue-like work" which "became immensely popular" and "remained without rival for a long time" (Willem F.H. Adelaar in Campbell & Grondona 2012: 13-4; Loukotka's name is mentioned in Campbell & Grondona allegedly as many as 508 times!).

[3] In 1998, somehow unexpectedly for many among the few interested in, vol. 4 (vi + 646) ("the first supplement to HAL", as Edward Vajda wrote in his review in *Language* 77/2 (2001: 360-3)) was released, with also two quite extensive grammatical outlines of Wai Wai (25-224) and Warekena (225-439).

[4] Living in Amazonas and Pará state border area on Nhamundá~Yamundá River (population 1242 in 2012, 942 and 1012 in 2010 quoted).

[5] Derbyshire's London University PhD dissertation *Hixkaryana Syntax*, dated 1979, rearranged and updated, was published in 1985; in its preface we read; "*In* the *six years* since the thesis was written, *there have been some notable advances*, both in the general research in linguistic typology and also *in published (or soon to be published) studies in Amazonian languages*" (vii; *italics* afm). Derbyshire's Hixkaryana texts, now a bibliopole and library rarity, appeared in print in 1965.

volume (Overall 2018) and the two volumes in focus of attention below, all within the range of this writer's sight, on the one hand, or Kalin 2011, 2014, on the other.

With the present review, after one-third of a century, LPos eagerly makes a comeback to the Amazonian languages (and, after two-thirds, to the beginnings of its interest in research on the indigenous South American languages) with the aim to introduce attractive recent results in the field to students and lecturers of general linguistics.

All four[6] vols. of Derbyshire & Pullum 1986-1998 totaled 2323 pp., being thus substantially larger in volume than the two vols. of Epps & Michael, here in focus (1476 pp. altogether) but, contrary to the latter, providing grammatical descriptions ("grammatical sketches") of (only?...) ten languages/ethnolects mentioned above (and in fn. 3) – one-third of the number of sketches included in the latter. On the other hand, the former offered various other articles (introductions, in vols. 1, 3, and 4 – typological[7] and comparative[8] studies, indices[9]), while Epps & Michael vols. have evidently been conceived in the first place to collect and introduce possibly all known isolated languages of Greater Amazonia (after all, presenting a language isolate is itself a result of comparative and typological research[10]); the editors and publishers declare in the one-page "Preface" (v) that "the subsequent volumes" are to address small and larger language families".

Understandably, the grammatical sketches of so many languages in only two volumes must have resulted in their length – they are (at times much) shorter than those included in Derbyshire & Pullum ((e.g. Warekena 215 pp., Wai Wai 200 pp., Piraha 125 pp.), but still the material offered is prognosticated to prove not only sufficient but optimal for "general" linguists and even more interested non-linguists)[11].

Following the 61-page long "front matter" is the core of the *Handbook* consisting of alphabetically arranged grammatical sketches referred to as <chapters> spreading over 1352 pages and two thick volumes in the following way:

– Vol. I – grammatical outlines of 13 languages: ***Aikanã*** (*~Mas*(*s*)*aká~ Kasupá~ Huari~Uari~Wari~Uapuruta~Waikorotá~Corumbiara~Kolumbiara~Tubarão*; by Hein van der Voort & Joshua Birchall; pp. 1-64; Brazil, Rodonia; approx. 250 users, 400 population), ***A'ingae*** (*~Cofán*; Rafael Fisher & Kees Hengeveld; 65-123; Columbia, Putumayo

[6] Cf. footnote 3.

[7] On certain typological features in Guajajara (Tupi-Guarani of Northeast Brazil) and on word order in Yagua (a sole Peba-Yaguan survivor of Northeast Peru) in vol. 1, and on "interclausal reference in Amahuaca" (Panoan of Peru and Brazil) in vol. 4.

[8] Two "Comparative Arawakan studies" in vol. 1, "A classification of Maipuran" (also Arawakan) in vol. 3, and a 130 pp. "Comparative study" of "Tupi-Guarani morphosyntax" (with five appendices) in vol. 4.

[9] "Cumulative" to vols. 1-3 in vol 3, and 1-4 in vol. 4.

[10] It is perhaps not out of place in just the present review to recall a section in LPos Loukotka 1950 (: 129-30) entitled... "Langues isolées" listing and localizing the following twelve tongues: Arikem~Uitate~ -Ahopovo, Gorgotoki, *Itonama*~Mačoto, *Kaničana*~Kanesi, *Kayuvava*, *Leko*~Lapalapa~Ateniano, Mašubi~Meken, Matanawi, *Mobimi*~Moyma, Parúborá, Toyeri ~Tuyoneiri~Arasairi~Huačipairi, *Yurakare*~Yuruxure~Kuči~Enete (some glottonyms *italicized* here purposefully).

[11] This reviewer was delighted to read (*ib.*) that "while linguists represent a primary audience for *this Handbook*, the work *is also intended to be accessible and useful to scholars and other interested readers from a range of disciplines and backgrounds*" (*italics* afm) – this is precisely what (and then – why he was) induced this author to produce the present text.

and Nariño, "severely endangered"; Ecuador, Sucumbíos, "definitely endangered"; respectively 379 and 600 for 2008**)**, ***Andoke*** (exoethnonym; Jon Landaburu; 125-72; Colombia, along Aduche river; pop. below 500, "linguistic validity low, Spanish in everyday use, most children do not speak Andoke"), ***A?iwa*** (15 ethnonyms listed, most with numerous variants; Christine Beier & Michael; 173-221; Peru, near Ecuador border; "two known rememberers 2008-10, minimally documented, now virtually extinct"), ***Betoi-Jirara*** **(***~Betoi*; Raoul Zamponi; 223-61; Venezuela, Apure; "a dialect cluster, today extinct"**)**, ***Cayubaba*** (Mily Crevels & Pieter Muysken; 263-300; northeast Bolivia; "handful of remembering elders in 2005, on the brink of extinction"), ***Chimane-Mosetén*** (*~Mosetenan*; Sandy Ritchie & Jeanette Sakel; 301-70; northern Bolivia; a dialect continuum, 13,500-16,000 users), ***Cholón*** (Astrid Alexander-Bakkerus & Kelsey Caitlyn Neely; 371--407; northern Peru; extinct, although speakers "living in the mountains" reported in 1996 or later), ***Guató*** (Kristina Balykova, Gustavo Godoy & Adair Pimentel Palácio; 409-40; Brazil; "on the edge of extinction"), ***Harakmbut*** (*~Harakmbet~ Hate*~mistakenly *Mashco*; An Van linden; 441-81; Peru, Cusco and Madre de Dios regions; six dialects/ethnic groups listed – one extinct, the other five "highly endangered, only a handful fluent speakers, if any, are left" in in five of the six ethnolects on the list; the entire population in or around 2007 quoted – 1,967[12]), ***Itonama*** (*~Sihnipadara*; Mily Crevels; 483-545; Bolivia, Beni provinces of Iténez and Mamoré; extinct[13]), ***Jodï*** (Jorge Emilio Rosés Labrada; 547-613; Venezuela, states Amazonas and Bolivar; 2011 census data indicate 854 speakers in the population of 982)[14], ***Kandozi-Chapra*** (Simon E. Overall; 615-57; northern Peruvian Amazon; population, according to the 2007 census, was 3,255 but "the current figure is likely to be higher"[15]), and "Index";

– **Vol. II** – grammatical sketches of 17 languages: ***Kanoé*** (*~Kanoê ~Kanoee ~Kanoä~ Kapishana~Kapišana~Kapixaná~Capixana*[16]; Laércio Nora Bacelar and H. van der Voort; 659-718; Brazil, southern Rondônia, Rio Omeré Indigenous Territory (*Terra Indígena*; three speakers (two monolinguals) mentioned in 2010, total pop. ca 325), ***Kwaza*** (H. van der Voort; 719-66; Brazil, southeastern Rondônia; "25 speakers out of an ethnic popula-

[12] "In general, young adults and speakers up to the age of fifty are bilingual in Harakmbut and Spanish, while speakers older than fifty are mainly monolingual in Harakmbut", revitalization efforts – didactic materials and governmental Bilingual Intercultural Education program signaled (443-4). The outline focuses on one dialect which has the highest number of speakers in the population of 1,043.

[13] "[...] there are no Itonama speakers left today" even if "the 2012 Census registered 16,158 people auto-identifying as Itonama, of which 1,249 claimed to speak the language as well [...]" (483).

[14] In his 2019 work, Rosés Labrada postulated, with "significant evidence" (549, fn. 2), Jodï to be a cognate of what he called a Jodï-Sáliban family. According to "Editor's note" (*ib.*), the text for the *Handbook* under concern here "is included in the isolates volume [...] because this chapter was commissioned prior to the publication of" the 2019 article and because the affiliation of Jodï has continued to generate some debate" (rearrangement – afm).

[15] Actually, "The Kandozi and Chapra are two indigenous groups" who "share the same culture and speak mutually intelligible varieties of the same language [...]" (615). It is by no means a unique ethnolinguistic situation or case (cf. e.g. various Mongolic groups, Karachays and Balkars, numerous peoples once united by what was known as Serbo-Croat(ian) or ... Swiss or Austrian speakers of German as their mother tongue, Russian-only-speaking Ukrainians or Byelorussians, etc.).

[16] Several more ethno-/glottonyms can be found in literature.

tion of about 45", 719, "the language survives in three separate families", 721), ***Máku*** (R. Zamponi and Chris Rogers; 767-806; Brazil, northwestern Roraima; extinct – allegedly in 1925 "little more than 50 speakers [... in] three communities", 767), ***Movima*** (Katharina Haude; 807-49; Bolivia, Beni; "in 2012 [...] spoken by approximately 500 adults" according to Movima's own count, "there are no first-language learners of Movima anymore", "severely endangered", 807), ***Muniche*** (some other names quoted in the text, 854; L. Michael, Stephanie Farmer, Gregory Finley, Kasrina Sullón Acosta, C. Beier, Alexandra Chanchari Icahuate, Donalia Icahuate Baneo, and Melchor Sinti Saita; 851--91; central Peruvian Amazonia; "almost extinct [...,] there are no longer any fluent speakers"[17]), ***Mỹky*** (*~Irantxe*; Bernat Bardagil; 893-937; Brazil, western Mato Grosso; "today, two distinct Indigenous communities speak different varieties of the Mỹky language: the Manoki and the Mỹky", 898[18]), ***Omurano*** (Zachary O'Hagan; 939-55; Peru, Loreto; "as of 2013, approximately 40 words and 15 short phrases were remembered collectively by fewer than 10 individuals [...] born between the early 1940s and 1980s [...]" their everyday language being Urarina, 939, "fluent Omurano speakers survived probably until the late 1990s or early 2000s", 944[19]), ***Pirahã*** (*~Apáitisí*; Raiane Salles[20]; 957-94; Brazil, Amazonas, Humaitá; "over 700" speakers "in ten villages", 957), ***Taushiro*** (Z. O'Hagan; 995-1027; Peru, Loreto; the last fluent speaker is introduced by his name, 995), ***Tinigua*** (Katherine Bolaños; 1029-75; eastern Colombia, Caquetá; again, "Today [i. e., at least 2019], the only known speaker of the language is [...] about 90 years old", 1029, although according to the national census results published in 2018 "a *Tanigua* group with 145 members and 28 speakers of the language is included (*Tinigua* is not mentioned). [...] the census does not specify the location of this group or how the data were collected", 1031), ***Trumai*** (Raquel Guirardello-Damian; 1077-1105; Brazil, Parque Indígena do Xingu, Terra Indígena Capoto-Jarina, and dispersed in the region; out of the population of "97 [...] there are 46 individuals who can speak Trumai, with different

[17] Among the very last (semi-)speakers were coauthors of the "Muniche" text (cf. p. 853). [...] the last fully-fluent speakers [...] were born between 1915 and 1925 and [...] the language was moribund by the early 1930s" (855).

[18] The described linguistic situation and its ethnic and historical contexts should draw particular attention of all interested or/and involved in studies of linguistic contact, language endangerment, preservation, extinction, planning, policies, extralinguistic influence, etc., etc. Necessarily short, the description offered can be indeed inspiring, cf.: "The 2014 [...] census counted 369 Manoki and 128 Mỹky. As a result of schooling, life in the Jesuit mission and extensive intermarriage with members of other Indigenous groups, language loss is severe among the Manoki. [...] in 2003 most Manoki under 50 were monolingual in Portuguese. [...] The language is much more vital among the Mỹky, a majority of whom are native speakers, although proficient knowledge of Portuguese is not uncommon. Both communities maintain frequent contacts and consider each other as part of the same people" (895).

[19] The Omurano speakers (but not only them) have been decimated by consecutive epidemics in the 20th century (cf. *ib.*).

[20] Not Daniel Everett... – the author of the almost 3.5 times more extensive "Pirahã" sketch in Derbyshire & Pullum 1986 and many other texts which made the language and its users world famous. In the 1986 (:200) sketch he classified the language as "a member of the Mura language family". Salles (957-9) writes that "There is a general concensus that Pirahã belongs to the Muran family, but the status of its extinct members as languages or dialects is unclear. [...] The most widely accepted contemporary classification, though, is that Mura is an isolate language and Pirahã is the last surviving of its numerous dialects"

degrees of proficiency [..., f]or the younger generations, Portuguese is now the main language of daily communication [...]", 1077), ***Urarina*** (Knut J. Olawsky; 1107-42; Peru, Loreto, Urarinas; "spoken by fewer than 3,000 people [...], in the more remote areas, Urarina is the first language of all speakers, including children", Olawsky mentions monolingual speakers in 2000-2005, 1107)[21], ***Wấnsöjöt*** (*~Puinave~Camaku~*; Adam J.R. Tallman, Cynthia Hansen, and Jesús Mario Girón; 1143-89; Colombia, Guainía, and Venezuela, Orinoco suroccidental venezolano; approximately 7,270 speakers, "an effort to introduce Wấnsöjöt in the schools, [...] courses [and] language revitalization programs" observed, 1145), ***Wao Terero*** (*~Waorani~Huaorani~ Huarani~Wao~Huao~Waotededo~ Wao Tirido~Auca~Auishiri~Awishiri~Sabela ~Ssabela*; Alexia Z. Fawcett; 1191-241; Ecuador, Pastaza and Napo, Waorani Ethnic Reserve; 1,766 speakers among 2,416 people who identify as Waorani, 1191), ***Warao*** (*~Warrau~Guarao~Guarauno* Andrés Romero-Figueroa and Konrad Rybka; 1243-82; northeastern Venezuela and adjacent parts of Guyana; pop. in Venezuela 32,000, shift to Spanish "conspicuous", "Warao monolinguals include all inhabitants of the isolated parts of the delta [Amacuro] , but only the elderly people elsewhere. In the eastern communities in Guyana [...] only a few elderly speakers remain [...], the [...] economic crisis in Venezuela drives the [...] Warao to migrate to Guyana and Brazil [...], 1245), ***Yaruro*** (*~Pumé*; Esteban Emilio Mosonyi and R. Zamponi; 1283-322; western Venezuela, Apume Llanos; "spoken actively and fluently by about 9,500 people"), ***Yurakaré*** (*~Yurújare ~Yuracare*; Rik van Gijn; 1323-52; central Bolivia, upper Mamoré River drainage area; 2-3,000 speakers, "today it is hard to find a child with active command of Yurakaré, although passive knowledge is still there", 1325), and "Index"[22].

The extensiveness of the front matter results from the decision of the designers of the *Handbook* to include the 45-page-long "Introduction" by the editors Epps and Michael – and there are perhaps reasons for doing so as it appears to concern the planned edition in its entirety, not only the two vols. under concern here, cf.: "The Handbook seeks to address these issues [signaled in two preceding paragraphs] by systematically compiling comprehensive, accessible grammatical overviews for *every* (*italics* afm) Amazonian language family and isolate for which adequate documentation exists" (xviii). "The burgeoning number of high-quality descriptive studies of Amazonian languages that have emerged over the past few decades" is mentioned (*ib.*) in this context, hence expect the appearance of (a rather large number of) consecutive volumes of the cover <HSK 44> in some near future.

The "Introduction" defines the term *Amazonian* used as an ethnonym and a glottonym to define what is and what will be included in or excluded from the entire Amazonian Handbook (xvii, xixf.), outlines (overwhelmingly tragic) histories of indigenous peoples

[21] Urarina "exhibits a range of unusual grammatical characteristics that are rare or absent in other languages" (1107); actually, from the point of view of "Western" linguists this can be said about perhaps all Amazonian languages.

[22] With few minor additions, the data provided in this longish paragraph have been intentionally restricted to what appeared in respective sketches-chapters and fastidiously selected with the focus on what the reviewer supposes to be most intriguing, fascinating for/to the possibly widest audience of potential users of such a *Handbook*.

and languages covered the term prior to ("a long view", xxif.) and following the "European invasion and beyond" (xxii-ix), and provides passages and sections of general linguistic and linguists' current interests, like language classification and diversity[23] (xxxf.), endangerment, shifts, revitalization efforts (xixf.), language contact (xxxivff.), unique typological features and phenomena ("Linguistic insights from Amazonia" section (xxxixff.) deserves special recommendation), outline of research (xxxviff.) from first documentations in the past, to present-day academic research, and "the view ahead" (xlif.)[24].

Since the two volumes described here concentrate on isolate languages, one more sentence from the "Introduction" deserves quotation: Some thirty of these genealogical units[25] are isolates" (xvii) – it allows the conclusion that the two *Handbook* volumes provide descriptions of *all of them* (see footnote 26, however).

The "Introduction" necessarily provides some space to language isolates as "the term "isolate" can be understood in a number of ways" (like no living relatives, no known relatives, poorly~unsatisfactorily attested sister languages with no remaining speakers; xviii) and a number of languages treated as isolated and therefore included in the *Language isolates* volumes were not treated as such in the past.

A map of a major (northern) part of Southern America with the localization (with points) of all the thirty described isolates is printed on p. xvi opposite the first page of the "Introduction" and systematically reprinted on the verso of the first page of each grammatical sketch but with toponyms of major geographical objects (cities, rivers), instead of glottonyms, marked, and with a smaller accompanying map of the area and/or region where the respective language is or was spoken or is or was associated with it and its speakers, placed beneath as projection from the point of the language localization on the larger map. Generally, the idea to organize the cartographical material in such a system, wisely conceived and consistently applied throughout the two volumes, could be only praised were it not the challenge the maps (or rather the size of characters printed on them) pose to this reviewer's elderly eyes (technically, the maps could easily be much more transparent with no or little cost to avoid this discriminatory treatment of potential users of the *Handbook*).

The two "Indexes", separate for each volume, seem open to doubt and vigilance: by far too short lists of entries being in turn too few and too general[26].

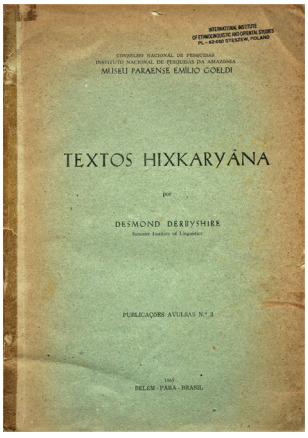
[23] Amazonia, comprehended territorially, ethno- and demographically, and linguistically (as far as ethnic and linguistic diversity is concerned) as defined in the "Introduction" is indeed comparable only to New Guinea and its neighborhood (cf. xxxf.).

[24] "In sum, the study of Amazonian languages is of urgent priority and enormous intellectual and humanistic value" (xlii).

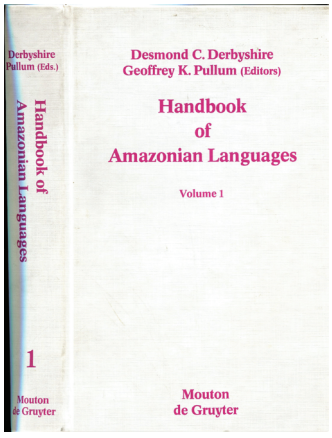
[25] i.e., "hundreds of Indigenous languages corresponding to some 100 distinct lineages" (*ib.*).

[26] e.g. browsing through the volumes, this writer spotted somewhere in the text mentioning the glottonym *Leko*; since it had been placed on Loukotka's 1950: 129 list of *langues isolées* but did not appear among the 30 sketches in the *Handbook*, this writer intended to revisit the spot using the indexes – in vain. Rationally – skimming through the almost 1,400 pages to look for one short word would be a nonsense. An experienced guessing game helped to find the suspicious fragment ("... a region of high linguistic diversity [...], with approximately 50 different languages pertaining to at least 18 different genealogical units [...] Seven language families... Besides these, 11 isolate languages are also spoken in that region [among them]: [...], Canichana, [...], Leko [...]", both on Loukotka list, cf. fn. 10, but untouched~neglected (?) in the *Handbook*...) in less

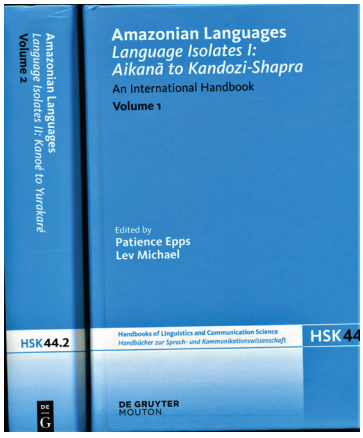
Data concerning demography (the entire population versus the population of speakers), linguistic situation (language attitudes, status, policies, etc.) of the communities and tongues involved, research (state of the art, current activity and prospects), etc., provided in the *Hand-book* clearly show that most of the languages described in the work are either seriously endangered or actually dead[27] and that rather few of them have been "adequately"[28] documented. Therefore, what this writer would like to point to as the first and foremost of what the De Gruyter Mouton HSK 44.1-2 publication offers and promises to offer is an extremely valu-able documentation of both the linguistic situation of Amazonia as defined by the authors, editors, and publishers as well as the thirty individual languages of the region described – and providing such documentation in view of the facts that up to 50% of the world's languages are endangered (or recently dead) and about 80% of them still remain "underdocumented" or completely undocumented doubtlessly is the most urgent task of linguistics and linguists. Virtually, a race against time which will be the factor determining the actual value, utility and presumed high rating of this new "handbook of Amazonian languages".



Derbyshire 1965



Derbyshire & Pullum 1986



Epp & Michael 2023

than ten minutes on page 897 (!) – this writer's modest and tiny success and satisfaction and an enormous failure of the *Handbook*! Without at least personal names, toponyms, ethnonyms, and glottonyms (there are plenty of each of them in the books) indexed, the edition can hardly serve as a reliable and handy reference work, unfortunately.

[27] Population smallness of a speech community is not necessarily a decisive factor in the processes of endangering or annihilating the community's tongue; the merciless killer is the abandonment of its intergenerational transmission (as social organisms all languages are doomed to change and finally die or undergo rebirth as new tongues with speakers finding the languages of their distant forefathers simply unintelligible).

[28] *i.e.* at least with existing comprehensive (practically it is "thick") two-way dictionaries, reference grammars, and representative text collections. Works – not incidentally, on Amazonian isolates described in the *Handbook* – like Olawsky 2006 (over 960 pp.!), van Gijn 2006 (370 pp.), Girón 2008 (496 pp.), but also less bulky ones like Pike & Saint 1988 (188 pp., "The texts" on pp. 105-68), could serve as examples here of what is needed.

References

Campbell, Lyle & Verónica Grondona (eds.). 2012. *The indigenous languages of South America: A comprehensive guide*. Berlin–Boston: De Gruyter Mouton.

Derbyshire, Desmond. 1965. *Textos hixkaryâna*. Belém Pará Brasil: Conselho Nacional de Pesquisas, Instituto Nacional de Pesquisas da Amazonia, Museu Paraense Emílio Goeldi.

Derbyshire, Desmond C. 1985. *Hixkaryana and linguistic typology*. Summer Institute of Linguistics & University of Texas at Arlington.

Derbyshire, Desmond C. & Geoffrey K. Pullum. 1979. Object initial languages. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 23(2). 1-34.

Derbyshire, Desmond C. & Geoffrey K. Pullum (eds.). 1986-1998. *Handbook of Amazonian languages.* Volume[s] 1 (1968), 2 (1990), 3 (1991). Berlin–New York (–Amsterdam [for vol. 1]): Mouton de Gruyter.

Dixon, R.M.W. (with the assistance of Alan R. Vogel). 2004. *The Jarawara language of Southern Amazonia*. Oxford: Oxford University Press.

Dixon, R.M.W. & Alexandra Y. Aikhenvald (eds.). 1999. *The Amazonian languages*. Cambridge: Cambridge University Press.

Girón, Jesus Mario. 2008. *Una grammatika del Wȧ́nsöjöt (Puinave)*. Utrecht: LOT [Landelijke Onderzoekschool Taalwetenschap].

Kalin, Laura Mennen. 2011. *Hixkaryana: The derivation of Object Verb Subject word order*. Los Angeles: University of California. (MA thesis.)

Kalin, Laura. 2014. The syntax of OVS word order in Hixkaryana. *Natural Language & Linguistic Theory* 32(4). 1089-104 (also *Coyote Papers* (University of Arizona) 20 (2012), 37-45).

LOS – *Linguistic and Oriental Studies from Poznań* (Adam Mickiewicz University, published 1991-2014; cf. fn. 2).

Loukotka, Čestmír. 1949. Sur quelques langues inconnues de l'Amérique du Sud. *Lingua Posnaniensis* 1. 53-82.

Loukotka, Čestmír. 1950. La parenté des langues du bassin de la Madeira[29]. *Lingua Posnaniensis* 2. 123-44.

Loukotka, Čestmír. 1951. [criticism of:] The languages of South American Indians by J. Alden Mason from [...] *Handbook of South American Indians*, vol. 6, [...], 1950. *Lingua Posnaniensis* 3. 366-75[30]

Loukotka, Čestmír. 1955. Les Indiens Botocudo et leur langue. *Lingua Posnaniensis* 5. 112-35[31].

Loukotka, Čestmir. 1968. *Classification of South American Indian languages*. Los Angeles: University of California Latin American Studies Center.

LPos – *Lingua Posnaniensis. Revue de Philologie Comparée et de Linguistique Générale* (Poznań), cf. fn. 2.

Olawsky, Knut J. 2006. *A grammar of Urarina* (Mouton Grammar Library 37). Berlin–New York: Mouton de Gruyter.

Overall, Simon E. 2018. *A grammar of Aguaruna (Iiniá Chicham)* (Mouton Grammar Library 68). Berlin–Boston: DeGruyter Mouton.

Pike, Evelyn G. & Rachel Saint (eds.). 1988. *Workpapers concerning Waorani discourse features*. Dallas: Summer Institute of Linguistics.

Pullum, Geoffrey K. 2008. Desmond C. Derbyshire (1924-2007). *Linguist List* 19.1 (Jan. 3, http://www.linguistlist.org/issues/19/19-1.html; reprinted in several typesetting versions in the Web).

Rosés Labrada & Jorge Emilio. 2019. Jodï-Sáliban: A linguistic family of the Northwest Amazon. *International Journal of American Linguistics* 85(3). 275-311 + 15 pp. "Online Supplement Material" (DOI: 10.1086/703238).

Van Gijn, Erik. 2006. *A grammar of Yurakaré*. Nijmegen: Radboud Universiteit. (Doctoral dissertation.)

[29] Rio Madeira, the largest tributary of the Amazon.

[30] The volume includes also Loukotka's "Supplementary Notes to the Classification of Australian Aboriginal Languages", pp. 135-57 + a large size map.

[31] LPos vol. 6 (1957) includes Loukotka's extensive "Classification des langues papoues", pp. 19-83 + a large size map; the volume publishes also "Notes on Huambisa Phonemics" by David Beasley and Kenneth L. Pike on pp. 1-8.

DOI: 10.14746/linpo.2024.66.1.5

Alexander Vovin, José Andrés Alonso de la Fuente, and Juha Janhunen (eds.) 2024. *The Tungusic languages*. Routledge Language Series. London and New York: Routledge. Pp. xxviii + 543

Alfred F. Majewicz

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

Die Zahl der in Sibirien befindlichen Tungusen
wird nach neuern statistischen Angaben [1854]
auf 35-40,000 Seelen beiderlei Geschlechts
geschätzt (Schiefner in Castrén 1856: 1)

Long awaited and long delayed (see p. xix), the above bibliographically described collectively authored monograph finally saw the light of the day in print with the © date of publication as indicated. Edited by three prominent specialists in the language area covered, the volume offers six (four following the front matter and two, preceding the "Index" (538-41) and concluding the core of the book) chapters-studies providing general information on the family in question in its entirety and fourteen chapters-descriptive sketches presenting individual languages, by sixteen authors.

Except for the detailed table of "Contents" (v-xii), the front matter includes an indispensable list of 112 "Tables" (xiii-xv), a list of "Contributors" (xvi-xvii; three of them deceased in 2020-2022), "Preface and Acknowledgements" (xviii-xix), "Technical Notes" (xx-xxi), a list of "Abbreviations" (xxii-xxiii), a "Chart of the Manchu Script" (xxiv-xxvi; tabularized but absent from the list afore mentioned), and a "Language Map" of the area involved (xxvii-xxiii). All this, together with the book-final index, make the volume very user-friendly indeed.

The initial chapters are consecutively devoted to: (**1**) introducing "Tungusic as a language family" (Janhunen, 1-18; here we have sections informing on "Ethnic nomenclature" (*i.a.* dealing with the ethno- and glottonyms <Manchu-Tungusic> ~ <Tungus-Manchu> but mainly clarifying Russian (old and new) and Chinese ethnonyms) but also such structurally organizing further chapters, like e.g. "Data and sources", "Typological profile", "Grammatical framework", "References and further reading" (see below); (**2**) "Early

Far-Eastern sources on Tungusic" (Vovin, 19-25); (**3**) "Early Western sources on Tungusic" (Alonso de la Fuente, 26-34), and (**4**) "Proto-Tungusic" (Janhunen, 35-75). The concluding "general" chapters deal with (**19**) "Sociolinguistic aspects of Tungusic" (Mamontova, 501-16; here e.g. "Demographic background" legislation, education, public sphere, ..., "Future prospects") and (**20**) "Tungusic in time and space" (Janhunen, 517-37; discussing i.a. the "Altaic"~"Macro-Altaic"~"Ural-Altaic hypotheses", "Altaicization" vs. "deAltaicization", Tungusic and Amuric (here interrelations with Gilyak~Nivkh), "the future of the Tungusic language family (which "looks bleak. The total number of Tungusic speakers in the world today is hardly more than 30,000 people, while the number of people transmitting the language to the next generation is probably in the hundreds, rather than thousands" (p. 532[1]); solely Solon is named as the language which "would seem to have the best chances of survival" [...], "somewhat surprisingly" (*ib.*), also for this reviewer).

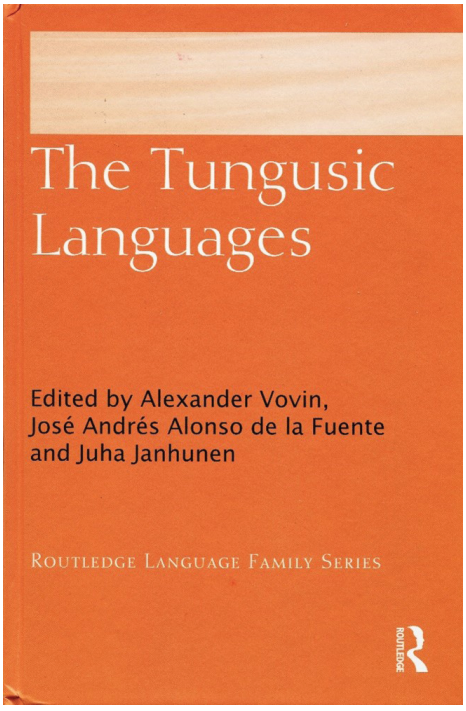
Chapters 5-18, in turn, have been dedicated each to one language: (**5**) extinct Jurchen (Kane & Miyake, 76-102); (**6**) Written Manchu (Vovin, 103-38); (**7**) Siberian Ewenki (Janhunen, 139-83); (**8**) Orochen (Chinese *Oroqen~Oroqon~ Elunchun* (*-zu* ethnonym 鄂伦春族, *-yu* glottonym 鄂伦春语; Whaley, 184-205); (**9**) Solon (Baek 白 尚 燁, 206-33); (**10**) Neghidal (Oskolskaya, 234-59), (**11**) Ewen (Malchukov, 260-93); (**12**) Oroch (Kazama 風間伸次郎, 294-325); (**13**) Udihe (~Udeghe~Udege~Ude~...; Perekhvalskaya & Janhunen, 326-63); (**14**) Nanai (~Nanay(an)~Chinese *Hezhe* (*-zu*, *-yu* 赫哲-族/-语; Kazama, 364-406); (**15**) Ulcha (~Olcha~Mangun~Santa(n); Kazama, 407-35); (**16**) Uilta (~Orok; Tsumagari & Yamada, 436-62); (**17**) Spoken Manchu (Zikmundová & Gao, 463-82, quotation: "According to information from 2017 no speakers use Manchu as a means of communication, but several individuals are capable of carrying a simple conversation, translating sentences from Chinese, or repeating short narratives they have memorized for the purpose of interviews[2]. [...] As of 2019, only one fluent speaker was believed to remain in the [Anhui] area", p. 464); (**18**) Sibe (~Xibe~Sibo~Xibo; Zikmundová, 483-500).

Apart from the competence of authors, worth emphasizing are the structure pattern of the individual language chapters (e.g., each, "Proto-Tungusic", "Tungusic as a language family", and the concluding chapter 20 included, either beginning with a section "Data and sources", or placing it second (six cases) or third (two, Siberian Ewenki and chapter 19, cases; *each chapter* ending with the bibliographical list of "references and further reading" mentioned; perhaps needless to stress, all language descriptions arranged in basically the same pattern allowing easy comparative reading and at the same time

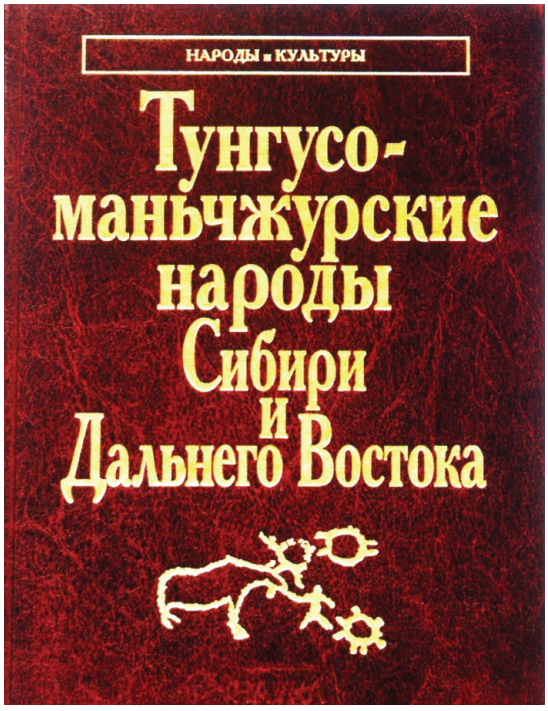
[1] Cf. the *motto* above, being the very first sentence in Schiefner's introduction to (and the first page in) Carstén 1856.

[2] It concerns Sanjiazi, "a relatively compact settlement, where Manchu was still used in everyday communication by c. 300 individuals in the 1960s" (*ibid.*). Interestingly, the process of dying out of the living (spoken) Manchu in its final phase resulted in the growing demand for, and proliferation of various handbooks of the tongue, readers, grammars, dictionaries of varying size and quality levels, including practical learning tools of the type "learn one phrase each day" (like He Rongwei's *Manyu 365 Ju* [365 Manchu phrases (like 'hello', 'how are you', 'everybody in family/home OK', 'goodbye', 'do you speak Manchu?', 'I know it a little', 'he can speak and write', 'when was the Manchu script created?', etc.] (何荣伟编 2009, [2]2012。满语365句 一天一句学满语。沈阳: 辽宁民族出版社)).

enriched with language specific material (e.g., “Taxonomic status”, “Dialectal division”, “Notational conventions” when appropriate), very similar to that implemented in Janhunen 2003[3]), “language-barrier-lessness” of source material and further- reading recommendations[4], and much updated state of the art in Tungusic studies than any other publication in the field (this reviewer noticed with great satisfaction references to Bronisław Piłsudski’s pioneering contributions to Tungusic studies with his Uilta~Orok and Ulchan data (but seemingly his Nanaian lexical material passed unnoticed in the respective chapter) made accessible to a wider audience only in 2011 as CWBP-4 as well as to Simonow-Kyalundzyuga Udeghe dictionary, now a rarity on the shelves of researchers[5].



Book under review (cover)



Missonova & Sirina 2022 (cover)

Thus, “thinking positively”, the delay in (as well as other, described in the preface, p. xix, serious perturbations and tragedies preceding) the publication of the book brought about unquestionable added value. It will take perhaps years to detect any possible (and

[3] Except for “Index”, much better and more than twice as long (13 pp.) as that in the Tungusic volume (6 pp.) under scrutiny here.

[4] Most of literature on Tungusic tongues has been published in Russian, Japanese, Chinese (in the case of China’s minority languages), and (understandably) in English but the authors quote or mention also works in German, and (rather rarities) Korean, French, Italian, Turkish, Polish, published in various countries.

[5] Unfortunately, a major part of the small edition stored perished in a fire and flood cataclysm in February 2003.

inevitable) flaws, errors, etc. Here, this reviewer would randomly point to only certain cases of overlooking (or deliberate omissions of) references[6] which, in his opinion, on varying reasons, deserve at least mentioning either in the respective sections of grammatical outlines of individual languages or in the "general information" chapters. A bit astonishing to this writer was the absence of a number of Boldyrev's bulky Ewenki dictionaries (only two Russian-Ewenki ones reported) 2000, 2009-10, 2010, 2013, ("less bulky") 2018, and his 2007 "morphology" or Myreyeva's 2004 Ewenki-Russian dictionary in the Ewenki chapter (144, 179-80), Robbek & Robbek 2005 Ewen-Russian dictionary in the Ewen chapter (262, 292), An's 2007 Chinese-Manchu dictionary (if its Manchu-Chinese predecessor is referred to (105, 134)), Aisin-Gioro Ulhicun's 1983 grammar (when her 1985 reader is listed, possibly also other Aisin-Gioro – Yingsheng and his 2004 miscellanea) or Nakajima al. 1999 and 1993-9 (characterized as "basic tools for Manchu linguistics, Chinese language, and Oriental history")[7] in the Written Manchu chapter, Yamamoto 1969 in the Spoken Manchu chapter (in spite of the title and languages involved[8]), Poniatowski 1923 (if Poniatowski 1966 and Simonov 1990 are (inexactly) referred to, cf. p. 328, 382, also Simonov 1991) in Nanai and Udihe chapters and Simonov & Dyachok 1995 in the Solon chapter; the reviewer could also expect to see mentioned in the delayed (hence – updated) publication works like e.g. a very reader-friendly Trofimova 2009 in Udihe chapter, Kazama 2003 for basics, or continuation of Kanchuga's autobiographical story in Udihe after 2006 (2016 volume is within this writer's sight); since "Orochen is among the least studied Tungusic languages" (185), this writer would draw (since "Orochen is among the least studied Tungusic language", 185) attention of the readers of the chapter on Orochen to Tulisow 1995 and Yin 1995.

The predecessor closest in design and character to what the book presented here offers was the extensive 180-page section on Tungus-Manchu languages ("Тунгусо маньчжурские языки") with an introduction by Orest Petrovich Sunik followed by eight grammatical sketches of Ewenki, Ewen, Neghidal, Nanai, Ulchan, Orok (Uilta) and Udihe in the fifth (and last) volume of the collective work entitled *Языки народов СССР* (pp. 53-232; see chapter 1 in the book under review, p. 5). It served for decades all those general linguists interested in, and greedy for information on, languages and linguistic structures substantially different from what they learned and studied at schools – and served well: minuscule editions of most other literature from the USSR were hardly available (and actually "too informative" for wider audience, while lan-

[6] The index proved to be useless for cross-checking.

[7] Perhaps also triglot Kurabayashi & Hurelbator 2006, 2008, if not other tetraglots and the famous pentaglot *Yùzhì wǔ tǐ qīngwén jiàn* 御製五體清文鑑 ("the high point of polyglot printing in the Qing dynasty", He 2018: 58). Actually, Vovin *did* write about these and such dictionaries (20-1) but did not list them in "References and further reading" (24-5) and the index, again, does not help.

[8] It is mentioned but only in the Sibe chapter (484, 500) in which for instance Lebedeva & Gorelova 1994 is not – so the confusion for linguists from outside the narrow field of Tungusology is mounting (again the editorial structure fails to help).

guage barrier played also its negative role). What we have now at our disposal is undoubtedly much better in every aspect[9].

*

Coincidentally, the release of the volume introduced here was published in almost the same time as another very attractive and also collective monograph (Missonova & Sirina 2022) on the peoples speaking the Tungusic languages with focus on cultural anthropology, economy, demography, etc., - the Ewenkis, Ewens, Neghidals, Uiltas, Nanais, Ulchas, Udihes, Oroches, and Tazes. Contemplating the two volumes complementing one the other, standing side by side on one shelf and neighboring 34 other impressive volumes from among those so far published in the «Peoples and cultures» series, must, however, remain in the sphere of dreams: *inter arma silent Musae, silent mores* – Academia cannot be a winner in the clash with operations classified as special.

References and abbreviations

(number of pages provided whenever considered meaningful and supportive)

Aisin-Gioro Ulhicun. 1983. *Manyu Yufa* [Manchu grammar Manchu]. Hohhot: Neimenggu Renmin Chubanshe. 爱新觉罗· 乌拉熙春编著 1983。满语语法。呼和浩特: 内蒙古人民出版社。

Aisin-Gioro Yingsheng. 2004. *Manyu zashi* [Manchu language miscellanea[10]]. Beijing: Xueyuan Chubanshe [Academic Press]. 1070 pp. 爱新觉罗·瀛生著 2004 满语杂识。北京: 学苑出版社。

An Shuangcheng (ed.). 2007. *Han-Man Da Cidian* (*Han-Man Duizhao*) [great Chinese-Manchu dictionary]. Shenyang: Liaoning Minzu Chubanshe. 1551 pp. 安双成主编 2007。漢滿 / 汉满大辞典。汉满对照。沈阳市: 辽宁民族出版社。

BGPU – Благовещенский госздарственный педагогический университет // Blagiveshchensk State Pedagogical University

Boldyrev, B[oris] V[asilyevich]. 2000. *Эвенкийско-русский словарь. Часть 1 А-П*, *Часть 2 Р-Я* // *Эвэды̄ лӯчады̄ турэ̄рук Ха̄н 1-2* [Ewenki-Russian dictionary, vols. 1-2; 988 pp.]. Novosibirsk: Издательство СО РАН Филиал «Гео».

Boldyrev, B[oris] V[asilyevich]. 2007. *Морфология эвенкийского языка* [morphology of Ewenki language]. Novosobirsk: «Nauka». Pp. 931 pp. !!!

Boldyrev, B[oris] V[asilyevich] 2018. *Словообразовательный словарь эвенкийского языка. Глагол // Эвэды турэн словообразовательнай турэрукин. Глагол* [Ewenki word building dictionary – verbs]. Novosibirsk: Nauka.

[9] Apart from the insufficient index, one more editorial decision should be pronounced wrong: the sole function of commas after the author(s)' name(s) in bibliographical listing arranged alphabetically should be to signal *the inversion* of name and surname in relation to what is to be found on the title page of sources referred to and/or listed: in a book with such accumulation of information on publications in languages in which family names consistently (as a rule) precede given names, inserting or not inserting the comma makes a huge difference.

[10] including parts on official (standard 规范语~written 书面语) language (here Chinese translation of Möllendorff's Manchu grammar appended), spoken language and dialects (满语口语及方言), Manchu-Chinese interactions (满语对汉语的影响), Manchu language "scavengings" (满语拾遗) and Chinese-Manchu glossary of common comparisons (汉满常用对照) – such a work, accessible in the Web, must not be left out uncommented in the book expected to "be informative for scholars and students [...] to obtain information on the Tungusic languages" (p. i; some information on Aisin Gioro Yingsheng and his two other publications can be found in the Spoken Manchu chapter, 466, 480).

Boldyrev, B[oris Vasilyevich] [&] G[ul'chera] V[akhobovna] Bykova [&] L[idiya] A[fanasyevna] Solovyova. 2013. *Словарь селемджинского говора эвенков Амурской области* // *Сэлэмдеды говор эвэды туруэрукин* [Selemdzha River Region Ewenki dialect dictionary]. *Часть* [vol.] *1* [A-Ӈ]. BGPU. 575 pp. (it is not clear whether vol. 2 has been published).

Boldyrev, B[oris] V[asilyevich] [&] G[ul'chera] V[akhobovna] Bykova [&] Galina I[vanovna] Varlamova [&] T[amara] Ye[gorovna] Andreyeva [&] R[aisa] Ye[gorovna] Malchakitova. 2009-10. *Словарь джелтулакского говора эвенков Амурской области* // *Дялтулакскай говор эвэды турэрукин. Элэкэсипты калтакан. Ге калтакан* [Dzheltulak River Region Ewenki dialect dictionary, vols. 1-2; (vol. 2: 2011 on cover)]. BGPU. 833 pp.

Boldyrev, B[oris Vasilyevich] [&] G[ul'chera] V[akhobovna] Bykova [&] Galina I[vanovna] Varlamova [&] L[idiya] K[onstantinovna] Senina. 2010 (2011 on cover). *Словарь зейского говора эвенков Амурской области* // *Зейскай говор эвэды туруэрукин* [Zeya River Region Ewenki dialect dictionary]. BGPU. 435 (+91 text specimens) pp.

Castrén, [Matthias]. 1856. *M. Alexander Castrén's Grundzüge einer tungusischen Sprachlehre nebst kurzem Wörterverzeichniss* [...] *herausgegeben von Anton Schiefner*. St. Petersburg: Kaiserliche Akademie der Wissenschaften.

CNEAS – Sendai: Tohoku University Center for Northeast Asian Studies // 仙台: 東北大学東北アジア研究センター

CWBP-4 2011. *The collected works of Bronisław Piłsudski, Volume 4. Materials for the study of Tungusic languages and folklore reconstructed, translated, and edited by Alfred F. Majewicz with the assistance of Larysa V. Ozoliņa* [&] *Mikhail D. Simonov* [&] *Tatyana Bulgakova* [&] *Elżbieta Majewicz* [&] *Tatyana P. Roon* [&] *Tomasz Wicherkiewicz* [&] *Werner Winter*. Berlin / Boston: de Gruyter Mouton. 1414 pp. "I. Materials for the study of the Uilta / Orok language and folklore", pp. 113-727, 739-45, 983, 1004, 1220-82, 1261, 1296-304. "II. Materials for the study of the Ulchan (Mangun) language and folklore", pp. 729-34, 736-8, 747-980, 1293-4. "III. Materials for the study of the Nanaian / Gold language and folklore", pp. 981-2, 986-1003, 1007-1214.

He Jiani. 2018. *From Empire to nation: The politics of language in Manchuria (1890-1911)*. University of Cambridge Faculty of Asian and Middle Eastern Studies. (Doctoral dissertation.)

IIEOS – International Institute of Ethnolinguistic and Oriental Studies, Stęszew, Poland.

ILCAA – Tokyo: University of Foreign Studies (Research) Institute for the Study of Languages and Cultures of Asia and Africa // 東京外国語大学アジア·アフリカ言語文化研究所

Janhunen, Juha (ed.) 2003. *The Mongolic languages* (Routledge Language Series). London and New York: Routledge.

Kanchuga, Aleksandr Aleksandrovich. 2016. *Багдисэ хокто телуунуни 8-2. Цаӣхи багдиӣ юӈдаӣни (2)* // *Автобиографическая повесть 8-2. Жизнь продолжается 2. Для чтения по удэгейскому языку* // *Autobiographical Story 8-2. Life goes on (2)*. An Udehe reader in the Udehe language with Russian translation, edited by Toshiro Tsumagari assisted by Kayo Kobayashi. Sapporo: Филологический Факультет Университета Хоккайдо // Hokkaido University Graduate School of Letters.

Kazama, Shinjiro. 2003. *Basic vocabulary (A) of Tungusic languages*. Suita, Osaka: Endangered Languages of the Pacific Rim.

Kurabayashi Hitoshi & Hurelbator (eds.). 2006. *«Gyosei manju mōko kanji sangō setsuon shimbunkan» mongorugo hairetsu taishō goi* // *Manchu-Mongolian-Chinese Triglot Dictionary of 1780 Arranged by Mongolian Words*. 653 pp. 栗林均・呼日勒巴特尔 編 2006。『御製滿珠蒙古漢字三合切音清文鑑』モンゴル語配列対照語彙。東北アジア研究センター叢書 第20号。

Kurabayashi Hitoshi & Hurelbator (eds.). 2008. «Gyosei manju mōko kanji sangō setsuon shimbunkan» manshūgo hairetsu taishō goi // Manchu-Mongolian-Chinese triglot dictionary of 1780 arranged by Manchu words. 574 pp. 栗林均・呼日勒巴特尔 編 2008。『御製滿珠蒙古漢字三合切音清文鑑』滿洲語配列対照語彙。東北アジア研究センター叢書 第30号。

Lebedeva, E[lena] P[avlovna] & L[idiya] M[ikhaylovna] Gorelova & V[asiliy] V[asilyevich] Radlov & Giovanni Stary. 1994. *Сиди Кур. Сибинская версия "Волшебного мертвеца" . Тексты в записи В. В. Радлова* // *Sidi Kur. A Sibe-Manchu version of the "Bewitched Corpse" cycle transcribed by V. V. Radlov; with transliteration and English introduction by Giovanni Stary*. Aetas Manjurica vol. 4. Wiesbaden: Harrassowitz Verlag. 422 pp.

LOS – *Linguistic and Oriental Studies from Poznań*

[Missonova, L[yudmila] I[vanovna] & A[nna] A[natolyevna] Sirina] (eds.). 2022. *Тунгусо-маньчжурские народы Сибири и Дальнего Востока. Эвенки. Эвены. Негидальцы. Уильта. Нанайцы. Ульчи. Удэгейцы. Орочи. Тазы* («Народы и культуры» series). Moskva: Наука. Pp. 1031.

Myreyeva, A[nna] N[ikolayevna]. 2004. *Эвенкийско-русский словарь.*// *Эвэды̄лӯчадыӈ тӯрэ̄рӯк* [Ewenki-Russian dictionary]. Novosobirsk: Nauka. 798 pp.

Nakajima Motoki [with] Imai Kenji [&] Takahashi Mariyō. 1993-9. *Den'nō shori Gyosei zōtei shinbunkan* // *Computational Analysis of Enlarged and Revised Manchu Dictionary Written by the Emperor* (*Yùzhì zēngdìng qīngwénjiàn* // *Han-i araha nonggime toktobuha manju gisun-i buleku bithe* [translated by He 2018:56-7 as *Imperially commissioned enlarged Manchu dictionary*]), vols. 1-9. ILCAA. 中嶋幹起編 今井健二·高橋まり代協力 1993。電脳処理 御製増訂清文鑑, 第一 ~ 九冊。

Nakajima Motoki [with] Imai Kenji [&] Takahashi Mariyō. 1999. *Shindai chūgokugo manshūgo jiten* [Qing Chinese-Manchu dictionary]. ILCAA. 1297 pp. 中嶋幹起編 今井健二· 高橋まり代協力 1999。清代中国語 満州語辞典。

Poniatowski, Stanisław. 1923. *Materials to the vocabulary of the Amur River Gold* // *(Materjały do słownika Goldów nadamurskich)*. Bibliotheca Universitatis Liberae Polonae. A, fasc. 10. Varsaviae: Universitatis Liberae Polonae.

Poniatowski, Stanisław. 1966. Dziennik wyprawy (do kraju Goldów i Oroczonów w 1914 roku) // Expedition Diary (to the Country of Golds and Orochons in 1914" [with the introduction of Barbara Heidenreich], followed by Franciszek Wokroj's Ludność nadamurska w świetle antropologicznych materiałów Stanisława Poniatowskiego // The Population of the River Amur in the Light of S. Poniatowski and Józef Glinka's Somatoskopia i somatotypologia (konstytucjonalna) grup nadamurskiej ludnosci Goldów i Oroczonów // Somatoscopy and Somatotypology (Constitutional) of the Groups of River Amur Population of Golds and Orochons". *Lud. Organe de la Société Polonaise d'Ethnologie* 50(1). 7-120 (including Goldski słownik i etnografia [Gold vocabulary and ethnography as recorded from an informant named Gendzu Aktanka from the village of Sakachialan], 95-9), 121-89, and 190-229.

[Robbek, V[asiliy] A[fanasyevich] & M[ariya Ye[gorovna] Robbek]. 2005. *Эвенско-русский словарь* [Ewen-Russian dictionary]. Novosibirsk: Nauka. 356 pp.

Schiefner, Anton 1856. Vorwort in Castrén 1856: v.

Simonov, M[ikhail] D[mitrievich]. 1991. *Удэгейская лексика в записях польского этнографа С. Понятовского* [Udihe words as recorded by Stanisław Poniatowski]. *Международная конференция Бронислав Пилсудский – исследователь народов Сахалина* [preceedings of the international conference "Bronisław Piłsudski – researcher of the peoples of Sakhalin", Sakhalin Regional Museum, Yuzhno-Sakhalinsk & Nogliki, Russia, October 31-November 2 (Yuzhno-Sakhalinsk)-November 3-4 (Nogliki)]. Vol. 2, 80-5. Yuzhno-Sakhalinsk: Сахалинский областной краеведческий музей [Sakhalin Regional Museum].

[Simonov, Mikhail Dmitrievich & Mikhail Timofyeyevich Dyachok]. 1995. Новые данные по лексике и грамматике солонского языка [new data on Solon lexicon and grammar]. *LOS* 2. 143-81.

Simonov, Mikhail Dmitrievich (compiler), Valentina Tunsyanovna Kyalundzyuga (principal informant), Alfred F. Majewicz (edited and published). 1998 (vols 1-2), 1999 (vol. 3). *A Kyalundzyuga-Simonov dictionary of the Udeghe (Udihe) language. Khor River Region dialect. Udeghe-Russian-Udeghe* // Кялундзюги-Симонова словарь удегейского языка удегейско-русско-удегейский (препринт). *Volume One* Vol. 1, pp. xx + 425. *Volume Two* 2, pp. xi+ 426-938, *Volume Three* xii + 939-1287). IIEOS Monograph Series 15/1-2-3. 1330 pp.

Trofimova, A[nna] O[legovna]. 2009. *Послелоги удэгейского языка* [Udihe postpositions~suffixes]. Novosibirsk: Сибирское отделение РАН, Институт филологии // Russian Academy of Sciences, Siberian Branch, Institute of Philology.

Tulisow, Jerzy.1995. New data on the Oroqen. *LOS* 2. 183-7.

Yamamoto Kengo. 1969. *Manshūgo kōgo kiso goishū* // *A classified dictionary of spoken Manchu with Manchu, English and Japanese indexes*. ILCAA. 山本謙吾著1969。満州語口語基礎語彙集。

Yin Tie-Chao. 1995. An Oroqen glossary and some of its related cultural aspects. *LOS* 2. 189-212.

DOI: 10.14746/linpo.2024.66.1.6

KIRSTEN REFSING (1948-2023)

This writer started, initially involuntarily, his interest in the Ainu and their language in mid 1970s, in relation to his queries concerning possible Polish contribution to global Japanese studies which resulted in "discovering" that practically all (not so numerous) publications of authors with Polish roots (Dybowski, Pilsudski, Radlinski, Sieroszewski) listed in renowned older (from Wenckstern 1895) international bibliographies of the discipline focused on the Ainu, particularly just their language and anthropology[1] (for query results see Majewicz 1979, 1993). That interest led to, and got solidified by, the rediscovery (actually for the fourth time) of Bronisław Piłsudski's 1902-1903 phonographic wax cylinder recordings of Ainu folklore and their preliminary inspection approximately around the time Refsing (as Taguchi) 1974 was published. Finding out that the book mentioned almost the same names (Dybowski, Piłsudski, Radliński (but also Ivan Kozyrevskiy ~ Ignacy Kosarzewski, cf. Golder 1914: 294-7)) was for this writer a motivating pleasure.

We both participated in a number of the same international conferences like at Warsaw (*Man in Contemporary Society and Culture of Japan*, Sept. 25-6 1981, cf. Refsing 1984a), Bochum (*Transcultural Understanding and Modern Japan*, Oct. 1-2 1981, cf. Refsing 1983a, Majewicz 1983) and Bonn (*History and Present State of Ainu Collections in Europe and European Studies on the Ainu*, June 24-7 1987, cf. Refsing 1993, Majewicz 1993) Universities, and in IBPC-1 (Hokkaido University, Sapporo and Nibutani *Ainu Bunka Shiryōkan*, Sept. 16-20 1985, cf. Refsing 1985, Majewicz 1985) and IBPC-3 (manggha Museum of Japanese Art and Technology in Kraków and Tatra Museum in Zakopane, Aug. 29-Sept. 7, 1999, cf. Refsing 2001, Majewicz 2001), and occasionally met in Hong Kong where Kirsten ruled over the Department of Japanese Studies at Hong Kong University[2] or in Japan.

In 1986 her most widely known and influential book – the grammar of one of southern Hokkaido Ainu dialects (actually, the idiolect of one of its allegedly two last female

[1] No wonder: "the Ainu race" became in the second half of the 19th century a long-lasting "sensation" not only in the academic world.

[2] 香港大學 in Pok Fu Lam 薄扶林 on Hong Kong Island 香港島 (not to be confused with the Chinese University of Hong Kong 香港中文大學 in Sha Tin 沙田).

speakers) was published in her native Denmark to soon meet with a warm reception. For instance, as early as 1987 the cultural anthropologist Bowles managed to print that "At last there is available to the English reader an intensive, in-depth analysis of the Ainu language, prepared by a competent Western scholar" (p. 671); similarly, the (also Denmark-born) linguist specializing in Japanese (especially historical) linguistics Takeuchi began (1988: 197) her review of the grammar with remarks that: "This attractively produced book presents the results of the author's field-work on one of the eastern Ainu dialects on Hokkaido, making available for the first time in English a firsthand, comprehensive account of that language". This writer, *and in this journal*, in turn concluded our review (1987, printed 1990) with equally enthusiastic welcoming words: "Refsing's book is both important and outstanding – and the only reliable reference grammar of the Ainu language so far published in any European language. It clearly surpasses in value any one of the Japanese grammars of Ainu [mentioned in the review] as well. [...] attractively edited, printed and bound [...], a splendid book in almost every aspect" (135), and an entire 6-page section (pp. 592-7) was devoted to Refsing's book by Dettmer in the third volume of his impressive four-volume "Ainu grammar" (1997) including an equally impressive overview chapter on "authors of writings on Ainu grammar and their works" (392-607, divided into three subchapters: "Ainu informants" 393-421, "Japanese scholars" 422-520, and "European and American scholars" 521-607) pronouncing that "the grammar represents a significant advance in European Ainu linguistics and in linguistic studies in general; the author took advantage of probably one of the last, after Batchelor[1], opportunities to collect, edit and publish previously completely unknown and unpublished material. Her description covers the entire thematic spectrum of Ainu linguistics"[2]. The situation "before Refsing" emerges from the "Introduction" to the "Appendix. A grammatical sketch of Ainu" in Patrie1982: 124-158: "Although a large amount concerning Ainu has been written in the Japanese language, the little that has appeared in English is totally inadequate and must not be relied on. [...] It is for this reason that this sketch has been included as an appendix to my book. I have translated, collated, and summarized the data from the various works of [Mashiho] Chiri, [Suzuko] Tamura, [Kyōsuke] Kindaichi, and [Shirō] Hattori. The various analyses presented in this chapter [sic!] are basically those of these authors" (*ib.*, 124). In other words, Refsing 1986 became a milestone in the history of Western studies on the Ainu.

[1] 1854-1944, contact with the Ainu 1878-1940.

[2] "Diese Grammatik bedeutet einen erheblichen Fortschritt in der europäischen Ainu-Sprachwissenschaft und darüber hinaus auch für dieses linguistische Gebiet allgemein; denn die Autorin hat, nach Batchelor, eine der letzten Möglichkeiten genutzt, bislang noch gänzlich unbekanntes und unpubliziertes Material zu sammeln, zu bearbeiten und zu veröffentlichen. Mit ihrer Darstellung behandelt sie die ganze Breite der Ainu-Sprachlehre, die Synopsis zeigt es" (596-597). In relation to our LPos review, Dettmer wrote: "Dieser Kritik des polnischen Sprachwissenschaftlers kann man in einigen Punkten ohne weiteres zustimmen, aber gewiß nicht in der Hindsicht, Refsings Opus sei höher zu bewerten als alle japanischen Ainu-Grammatiken, wenngleich die Verfasserin die Wörter anders klassifiezert" (*ib.*, 595).

Kirsten participated in two of the so far four International Conferences on Bronisław Piłsudski and His Legacy in Sapporo and Nibutani (IBPC-1, 1985; left, cf. Refsing 1985) and in Kraków (IBPC-3, 1999; right, cf. Refsing 2001) and Zakopane

That history of exploration, collection, and research by Westerners turned out to be ages-long and astonishingly rich in travelogues and records of both material and intangible culture as well as language data from virtually all regions inhabited by the Ainu, and since the Ainu language is now actually extinct and the people underwent a sweeping acculturation within the period between ca mid 19th century – ca 1930s, such testimonies, collections, and records became invaluable for further investigations in many disciplines. Access to those written sources, especially those of earlier dates, however, was very limited not only because of their age but also because most of them were printed in limited editions in remote places with turbulent and tragic past. The situation changed dramatically with the publication of Refsing's 25-volume 8094 pp. anthology of reprints of 165 such works by about 84 authors (Dybowski, Radliński 1891 and Piłsudski 1912 included) in four series (see EEWCR, EEWCT, EEWI, EEWL) entitled *The Ainu Library Collection*, each series preceded with Refsing's competent extensive introduction in vol. 1 (cf. Refsing e.g. 1996, 2000).

With such experience and published research results she was predestined to substantially contribute to at least the Ainu volume (Bugaeva 2022) in the HJLL series, and at least particularly authoring the planned chapter on "European /Western" records of the language as this writer recommended. Yet, it was she who recommended otherwise. Apart

from perhaps the innermost circle of her acquaintances, hardly anyone well acquainted with her academic output could perceive her denial as a signal of something getting wrong, judging rather, with full understanding, that she was too busy to squeeze any consecutive obligation. Kirsten Refsing was really discreet, only now scanty but positively moving memories in Danish from her last place in this world – Plejehjemmet Slottet in Nørrebro, the palace in which she could enjoy the autumn of her life – slowly make her personality more complete.

Born and educated in Fredericia in south eastern Jutland (Region Syddanmark), Kirsten graduated (MA) in Japanese from the University of Copenhagen in 1977 and it was her 1986 grammar that secured for her the PhD degree from the same University in which, after 1984-90 years spent as assistant professor at Aarhus University, she served as Associate Professor and Head of Department at the East Asian Institute (Østasiatisk Institut) and in 2006-11 as Dean of the Faculty of Humanities. 1995-2006 were her Hong Kong years mentioned above.

Although her publications make their author an unquestionable authority on the Ainu and one of the most important names in the entire bulk of Ainu (thence, of course, also Japanese) studies (cf. selected titles in the bibliography appended), and at least one of aids for Japanese text translation coauthored by her (with Lundquist, 2008) had worldwide circulation, she was also engaged in translation practice (cf. e.g. Kamo-no Chōmei 鴨長明's 1212 *Hōjōki* <方丈記> into Danish published as *Den vemodige eneboer* 'the melancholy recluse', 2020), editorship of collective volumes (like *Ideology and Practice in Modern Japan*, co-editor Roger Goodman, 1992), and writing... novels in English (at least two – *Eliza, A Missionary Wife*, 2018, and *Changes in the Shadows*, 2020). Much about *this* Kirsten Refsing the author of the present text learned, however, only in the course of composing it in March 2024. The distressing news reached him in late January 2024. Too many persons leave us too early...

†

Dr.phil.

Kirsten Refsing

* 14. september 1948 † 22. september 2023

er taget videre fra dette liv.

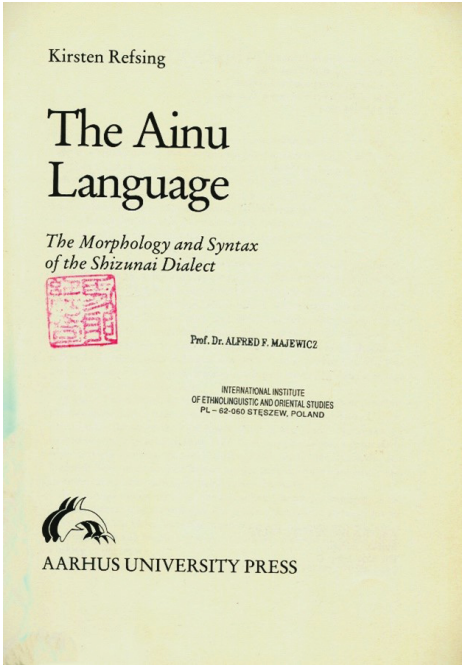
Bisættelsen finder sted fra De Gamles Bys Kirke
fredag den 6. oktober kl. 13.00

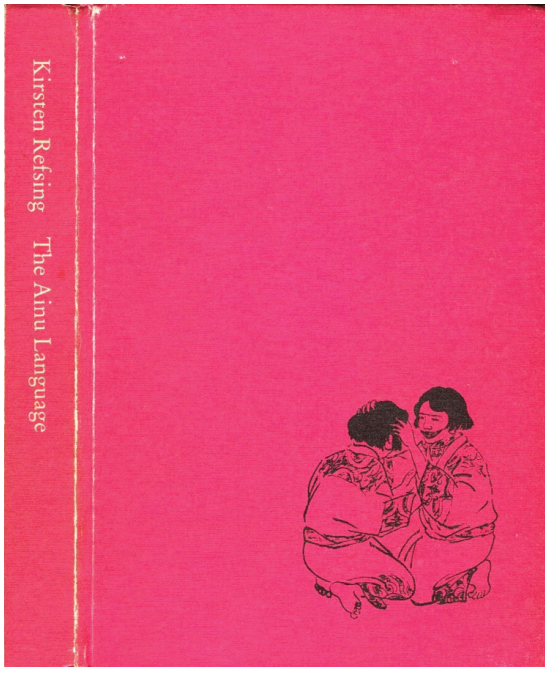
Til bisættelsen kan medbringes en blomst/gren/lille genstand til fælles pyntning af kisten.

Familien

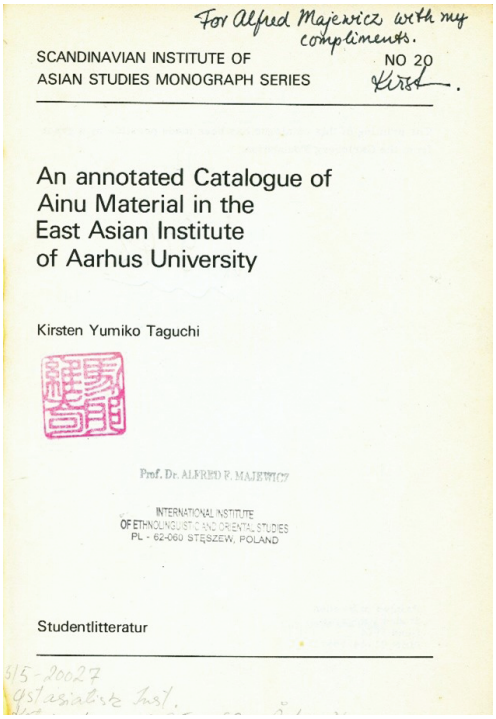
References and bibliography

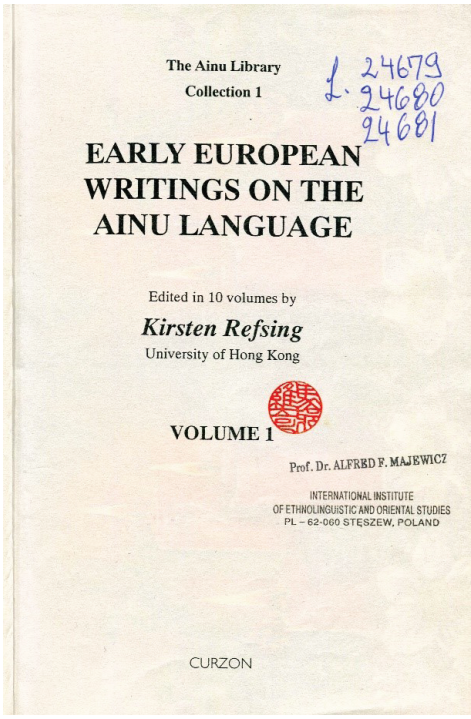
Bowles, Gordon T[owsend]. 1986. Review of Refsing 1987. *The Journal of Asian Studies* 46(3). 671-672.
Bugaeva, Anna 2022. *Handbook of the Ainu language*. *Volume 12* in the series *Handbooks of Japanese Language and Linguistics* [HJLL]. Boston–Berlin: De Gruyter Mouton.
CWBP-2 – *The collected works of Bronisław Piłsudski*, *Volume 2, Materials for the study of the Ainu language and folklore (Cracow 1912)*, edited by Alfred F. Majewicz. Berlin–New York: Mouton de Gruyter. 1998.
Dettmer, Hans Adalbert. 1997. *Ainu-Grammatik. Teil II: Erläuterungen und Register A*. Wiesbaden: Harrassowitz Verlag.
Dybowski – see Radliński 1891 and Murayama 1971.
ECIS – Executive Committee of the International Symposium [Toshimitsu Asakura et al.] (eds.) 1985. *Proceedings of the International Symposium on B. Piłsudski's Phonographic Records and the Ainu Culture, September 16-20, 1985, Hokkaido University*. Sapporo: Hokkaido University.
EEWCR – Kirsten Refsing (ed.). 2002. *Early European writings on Ainu culture 2. Religion and folklore*. Vols. 1-5. (The Ainu Library Collection 4). London: Routlege Curzon & Tokyo: Edition Synapse.
EEWCT – Kirsten Refsing (ed.). 2000. *Early European writings on Ainu culture 1. Travelogues and descriptions*. Vols. 1-5. (The Ainu Library Collection 3). Richmond: Curzon & Tokyo: Edition Synapse.
EEWI – Kirsten Refsing (ed.). 1998. *Origins of the Ainu language: The Indo-European controversy*. Vols. 1-5. (The Ainu Library Collection 2). Richmond: Curzon.
EEWL – Kirsten Refsing (ed.). 1996. *Early European writings on the Ainu language*. Vols. 1-10. (The Ainu Library Collection 1). Richmond: Curzon.
Golder, F[rank] A[lfred]. 1914. *Russian expansion on the Pacific 1641-1850*. Cleveland: Arthur H. Clark Company.
IBPC – International Conference on Bronisław Piłsudski and His Scholarly Legacy.
IIEOS – International Institute of Ethnolinguistic and Oriental Studies, Stęszew, Poland.





Refsing 1986: the first modern Western-language 301-page grammar of Ainu; title page, back, and cover





Refsing 1974 (left) and EEWL vol. 1 title pages

Johansen, Rebecca Veinø 2022-3. [randomly found e-journalism from Danish Web].

Kracht, Klaus & Helmut Morsbach (eds.). 1983. *Transcultural understanding and modern Japan: Fourth Symposium Sponsored by Tôkai University, held at the Ruhr University. Bochum, West Germany, October 1981*. Bochum: Studienverlag Dr. Norbert Brockmeyer.

Kreiner, Josef (ed.). 1993. *European studies on Ainu language and culture.* München: iudicium Verlag.

LPos – *Lingua Posnaniensis. Revue de Linguistique Comparée et Generale.*

Majewicz, Alfred F. 1976-9. *The importance of the Polish contribution to the study of the Ainu and their language*. PIJUAM 1976. Pp. 69. Mimeographed, later partially published as [16] "The importance of the Polish contribution to the study of the Ainu and their language". LPos 22 (1979), 93-122 and "On B. Piłsudski's unpublished Ainu material". *Hoppō Bunka Kenkyū* 11 (1977), 83-94 (reproduced in Asakura T. & Ifukube T. (eds.) 1986. *Piusutsuki rokuon rōkan kenkyū-no ayumi* [documentation of the research on B. Piłsudski's phonographic records]. Sapporo: Hokkaidō Daigaku Ōyō Denki Kenkyūjo, pp. 168-79) with Erratum in LPos 22 (1980), 122, and „Pōrando-no ainu kenkyū, Karafuto chōsa-nado furui rekishi" [Polish studies on the Ainu, the old story of Sakhalin data]. *Mainichi Shimbun – yūkan* (Osaka) Feb. 23, 1977 (the latter also mimeographed as *Pōrando-ni okeru ainu kenkyū-no rekishi-ni tsuite // On the history of Ainu studies in Poland.* PIJUAM WP 3. 1982. Pp. 24).

・北方文化研究　。

・朝倉利光・伊福部達編集1986。ピウスツキ録音蝋管研究の歩み。札幌:
北海道大学応用電気研究所。

・ポーランドのアイヌ研究, 樺太調査など古い歴史。毎日新聞 夕刊。

・ポーランドに於けるアイヌ研究の歴史について。

Majewicz, Alfred F. 1983. Linguistics and the problem of the resident Ainu population in Southern Kamchatka. In: Kracht & Morsbach 1983: 132-38.

Majewicz, Alfred F. 1985. Vicissitudes of B. Piłsudski's lexicological collections. ECIS, 173-81.

Majewicz, Alfred F. 1987. Review of Refsing 1986. LPos 31 (1990), 129-35.

Majewicz, Alfred F. 1993. On Poles collecting data related to the anthropology and material culture of the Ainu. In: Kreiner 1993: 127-36.

Majewicz, Alfred F. 2001. Bronisław Piłsudski, Adomas Varnas and the "Lithuanian Crosses". In: Majewicz & Wicherkiewicz 2001: 75-81 + photos 6-19 (separate plates).

Majewicz, Alfred F. 2022. 4. Ainu language Western records. In: Bugaeva 2022: 99-145.

Majewicz, Alfred F. 2023. A new vision and new frontiers of Japanese linguistics. <Ainu Language Western Sources> in the HJLL 12 Ainu Volume [an online presentation and lecture at the International Academic Conference *Practicing Japan – 35 Years of Japanese Studies in Poznań and Kraków* March 24-26, 2022]. Also distributed as: *A new vision and new frontiers of Japanese linguistics. <Ainu Language Western Sources> in the HJLL 12 Ainu Volume*. IIEOS Preprint 60 (2021) and published as "A new vision and new frontiers of Japanese linguistics. Ainu language Western sources in the handbooks of Japanese language and linguistics (12) Ainu Volume". *Silva Iaponicarum* 69 summer // 日林第六十九夏 (2023), 55-116.

Majewicz, Alftred F. & Tomasz Wicherkiewicz (eds.). 2001. *Bronisław Piłsudski and Futabatei Shimei – an excellent charter in the history of Polish-Japanese relations: Materials of the Third International Conference on Bronisław Piłsudski and His Scholarly Heritage, Kraków-Zakopane 29/8-7/9 1999*. *Linguistic and Oriental Studies from Poznań* Monograph Supplement 7. Poznań: Adam Mickiewicz University Press.

Murayama Shichirō. 1971. *Kita Chishima ainugo – bunkengakuteki kenkyū* [philological studies on Northern Kuril Ainu language]. Tōkyō: Yoshikawa Kōbunkan. 村山七郎。北千島アイヌ語 -- 文献学的研究。東京: 吉川弘文館。

Patrie, James. 1982. *The genetic relationship of the Ainu language*. [Honolulu]: University Press of Hawai'i.

PIJUAM – Poznań: Adam Mickiewicz University Institute of Linguistics.

Piłsudski, Bronisław. 1912. *Materials for the study of the Ainu language and folklore* collected and prepared for publication by Bronisław Piłsudski, edited under the supervision of J[an] Rozwadowski, PhD. Cracow: Imperial Academy of Sciences (Spasowicz Fund) „Spółka Wydawnicza Polska". Reprinted several times, also in CWBP-2, 1-272; Japanese and Russian translations available.

Radliński, Ignacy. 1891. *Słownik narzecza Ainów, zamieszkujących wyspę Szumszu w łańcuchu Kurylskim przy Kamczatce ze zbiorów Prof. B. Dybowskiego* [dictionary of the ethnolect of the Ainu inhabiting the Shumshu Island in the Kuril chain nearest to Kamchatka Peninsula from records taken by Professor Benedykt Dybowski]. Kraków: Akademia Umiejętności.

[Refsing], Kirsten Yumiko Taguchi. 1974. *An annotated catalogue of Ainu material in the East Asian Institute of Aarhus University*. Lund: Studentlitteratur.

Refsing, Kirsten. 1983. Ainu language. In: *Kodansha encyclopedia of Japan*, vol. 1. Tokyo: Kodansha Ltd. Pp. 36-37.

Refsing, Kirsten. 1983a. The Ainu minority in Japan. In: Kracht & Morsbach 1983: 121-31. 121-31.

Refsing, Kirsten. 1984. Verb auxiliaries in the Shizunai dialect of Ainu. *The Annual Report of the Regional Research Institute, Asahikawa University* 6, 61-86.

Refsing, Kirsten. 1984a. Personal affixation in the Shizunai dialect of Ainu. In: Mikołaj Melanowicz (ed.), *Man and society in Japan today*. Warsaw University Press. Pp. 59-70.

Refsing, K[irsten]. 1985. Sentence embedding in the Ainu language. In: ECIS 143-150.

Refsing, Kirsten. 1986. *The Ainu language: The morphology and syntax of the Shizunai dialect*. Århus: Aarhus University Press.

Refsing, Kirsten. 1993. The Ainu concept of time as expressed through language. In: Kreiner 1993: 91-101. 91-101.

Refsing, Kirsten. 1994, [²2011]. Tense, aspect and actionality in the Ainu language. In: Carl Bache & Hans Basbøll & Carl-Erik Lindberg (eds.), *Tense, aspect and action: Empirical and theoretical contributions to language typology*. Berlin–New York: Mouton de Gruyter. 311-22.

Refsing, Kirsten. 1996. Introduction [to EEWL 1-10]. EEWL 1, 1-35.

Refsing, Kirsten. 2000. Introduction [to EEWCT 1-5]. EEWL 1, 1-107.

Refsing, Kirsten. 2000a. Lost Aryans? John Batchelor and the colonization of the Ainu language. *interventions* 2(1), 21-34.

Refsing, Kirsten. 2001. Early travelers among the Ainu. In: Majewicz & Wicherkiewicz 2001: 159-71.

Refsing, Kirsten. 2014. From collecting words to writing grammars: A brief history of Ainu linguistics. In: Mark J. Hudson & ann-elise lewallen & Mark K. Watson (eds.), 2014. *Beyond Ainu studies: Changing academic and public perspectives.* Honolulu: University of Hawai'i Press. 185-99.

Refsing, Kirsten & Lita Lundquist. 2008, 2009 [3 editions]. *Translating Japanese texts*. Museum Tusculanum Press.

Refsing, Kirsten & Lita Lundquist. 2009. *Translating Japanese texts // Nihombun-no hon'yaku*. Copenhagen: Museum Tusculanum Press // Folrag. 日本文の翻訳

Taguchi – see [Refsing] 1974.

Takeuchi, Lone. 1988. Review of Refsing 1986. *Bulletin of the School of Oriental and African Studies* 51(1). 177-8.

Wenckstern, Fr[iedrich] von, 1895. *A bibliography of the Japanese Empire being a classified list of all books, essays and maps in European languages relating to Dai Nihon* [...]. Leiden: Brill.

Alfred F. Majewicz
IIEOS