



Functional Domains, Functions, and the Notion of Complexity: The Systems of Reference

Zygmunt Frajzyngier*

Department of Linguistics, University of Colorado Boulder, Boulder, CO, United States

The present study addresses the issues of (1) how to define complexity in the study of functions, (2) how to measure complexity in the study of functions, and (3) the benefits of the notion of semantic complexity in the analysis of language. This argues for a metric of complexity narrowed to single domains, something that has been already mentioned in some other studies. Such measures of complexity can then point to areas of further studies, both synchronic and diachronic. Two metrics of complexity are proposed: The first one involves the number of functions encoded in the given domain. The second is the number of functions that the speaker needs to take into consideration in realizing the functions encoded in the given domain. The argumentation for the proposed approach to complexity is based on cross-linguistic examination of the systems of reference of languages belonging to different families. The implication of this study is that the complexity of functional domains is the fundamental motivation of the complexity of the formal means of coding.

OPEN ACCESS

Edited by:

Marcin Maria Kilarski,
Adam Mickiewicz University, Poland

Reviewed by:

Piotr Gąsiorowski,
Adam Mickiewicz University, Poland
Nina Pawlak,
University of Warsaw, Poland

*Correspondence:

Zygmunt Frajzyngier
zygmunt.frajzyngier@colorado.edu

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 27 October 2020

Accepted: 24 March 2021

Published: 05 May 2021

Citation:

Frajzyngier Z (2021) Functional Domains, Functions, and the Notion of Complexity: The Systems of Reference.
Front. Commun. 6:622105.
doi: 10.3389/fcomm.2021.622105

Keywords: reference system, measuring complexity, typology, functional domains, coding means

INTRODUCTION

For more than a 100 years, the study of language complexity has had complete languages in its scope (McWhorter, 2001a,b, 2009; Sampson, 2009; Newmeyer and Joseph, 2012; Dixon, 2016: Chapter 6, 125–146). For an excellent review of the current approaches to linguistic complexity, see Dahl (2004). Dahl offers a study of changes in linguistic complexity with the focus on morphology. Older studies, many of which focused on morphology, asked the question of which languages are more complex and which are less complex. Modern studies examine relations between complexity and a given linguistic theory, language change (Sampson, 2009), language contact, the nature of creoles and pidgins (McWhorter, 2009), first- and second-language acquisition, and the relationship of complexity to non-linguistic factors, such as the size of population, physical environment of the speakers, and cultural norms. All of these are legitimate areas of study justified by the discussions they engender. However, other than the important question of the relationship between first- and second language acquisition, it is not clear what are the heuristic advantages of whole-language complexity studies, apart from the study of complexity itself.

The term “complexity” in the present study refers to the number of functions the speaker **must** include when forming a predication in a given domain. The larger the number of functions to be processed, the larger is the complexity in the given domain. The present study addresses the issue of how complexity could serve in linguistic analysis, namely the relationship between coding means available in the language and the complexity of functions. The study argues for

two types of metrics of complexity. Each metric of complexity should be narrowed to a specific single domain, something that has already been mentioned in some other studies. The first metric involves the number of functions encoded in the given domain. The constituents of each function may have connections with other domains, and the speaker needs to take those connections into consideration. Those connections constitute the second metric that the speaker needs to take into consideration in realizing the functions encoded in the given domain. Such measures of complexity can then point to areas of further studies, both synchronic and diachronic, e.g., the emergence and growth of complexity; the decrease and loss of complexity; and the consequent emergence and loss of functions, and possibly forms, or changes in function (Frajzyngier and Butters, 2020).

The theoretical framework for the present study is as follows: Every grammatical system encodes a finite semantic structure, which is comprised of functional domains. Each functional domain is comprised of functions. All functions within a given functional domain share a single feature that defines the domain, and all functions within the domain must differ from each other with respect to a single feature that defines the function. The determination of functions and features is based on analysis of the formal means of coding within the given language, including prosodic and phonological means, lexical categories, inflectional morphology on all lexical categories, linear orders, and deployment of lexical items to code grammatical functions, e.g., serial verb constructions, and possibly others. Languages differ in the number of functional domains they encode in the grammatical systems and in the internal structures of functional domains. The uniqueness of this approach rests on the fact that the determination of the function or meaning is based on the relationship with other functions within the same functional domain rather than on inferences about reality resulting from the use of certain forms (Frajzyngier and Shay, 2003; Frajzyngier with Shay, 2016; Frajzyngier and Butters, 2020).

Complexity, when confined to the study of the internal structures of similar functional domains, can be formulated in terms of the number of functions coded in a given domain. The heuristic advantage of such a study of complexity is that it forces the researcher to state explicitly whether a given function is a member of a functional domain consisting of two functions, three functions, four functions, or more. The description of a function

is crucially dependent on the contrast with other functions within the same domain.

The present work is a cross-linguistic illustration of a study of complexity within the systems of reference. Systems of reference have an impact on the related domain of coding relations between the verbal predicate and participants in the proposition and on the forms of utterances in the language.

REFERENCE SYSTEMS

The traditional meaning of the term “reference” is the “relationship between a part of utterance and an individual or a set of individuals that it identifies” (Matthews, 1997: 312). For some contemporary approaches to reference in philosophy, psychology and linguistics, see Gundel and Abbott (2019). In the present study, the term “reference system” designates all functions within the grammatical system of the given language that indicate (a) **whether** the listener should identify the participants in the proposition and, if so, (b) **how** they should identify the participants. The coding means within the system of reference may include: deployment of a noun phrase; the absence of a noun phrase in a position where it can be deployed; many types of pronouns, with each type having a different function; gender and classification systems and their indexing on a variety of lexical categories such as nouns, adjectives, numerals, and demonstratives; markers of agreement used on the verb and other lexical categories, including prepositions, demonstratives, determiners and articles, linear orders, complementizers, and conjunctions; inflectional markers coding same or switch reference; and a variety of prosodic means including tone, intonation, stress, and pauses. Each of these coding means has a function that is defined by its interaction with other functions within the reference system.

The proposed approach postulates that the relationship between form and function or meaning, including reference, is not direct but rather is mediated by the intermediary relationship between the functions within a given domain. One function differs from other functions by just one feature. Here is an illustration of these two principles in the domain of reference. If a language e.g., Mupun (West Chadic, Frajzyngier, 1993) codes the category “previous mention,” this creates a binary functional distinction between previous mention and lack of previous mention. The speaker therefore has to indicate whether the noun has been previously mentioned or not. In a language that does not code the function of previous mention the speaker does not have to address this function. Similarly, if the grammatical system encodes logophoricity, the speaker has to indicate whether the participants in the complement clause are coreferential or non-coreferential with the participants of the matrix clause (Frajzyngier, 1985, 1993). For the elaboration of the theoretical approach taken in this study, see Frajzyngier and Shay (2003), Frajzyngier with Shay (2016), and Frajzyngier and Butters (2020).

The research on the reference systems indicates that the same sets of forms across languages may carry opposite values within the same functional domain. Here are a few examples: The

Abbreviations: 1, 1st person; 2, 2nd person; 3, 3rd person; ACC, accusative; ADJ, adjective; ANAPH, anaphora; ASSC, associative; COM, comment marker; COMP, complementizer; CONJ, conjunction; CONJ:a, unexpected follow-up conjunction used to conjoin clauses, and utterances; CONJ:i, coordinating conjunction used to conjoin nouns, clauses, and utterances; D, dependent (aspect); DAT, dative; DED, deduced reference; DEM, demonstrative; DIM, diminutive; DU, dual; EE, end of event marker; EXCL, exclusive; F, feminine; F., Fula (Fulfulde); FUT, future; GEN, genitive relationship (not necessarily genitive case); GO, goal orientation; HAB, habitual; IMP, imperative; INCL, inclusive; INF, infinitive; INS, instrumental; INTENS, intensifier; IPFV, imperfective; LOC, locative; M, masculine; N, neuter; NEG, negative; NKJP, Narodowy Korpus Języka Polskiego; NOM, nominative; OPT, optative; PASS, passive; PFV, perfective; PL, plural; POL, polite request; POS, point-of-view of subject; POSS, possessive; PRED, predicator; PREP, preposition; PRES, present; PRS, presentative; PST, past; PUNCT, punctual; REFL, reflexive; REL, relative marker; REM, remote; SG, singular; STAT, stative; TOP, topicalizer; TP,thetic predication.

deployment of subject pronouns in Polish (Slavic) indicates that the subject of the clause is in focus or is different from the subject of the immediately preceding clause. Subject pronouns in English carry no value as to whether their referents are the same as, or different from, the subjects of the preceding clause (Frajzyngier, 1997). Coding of the third-person subject on the verb in Polish indicates that the subject of the clause is the same as the preceding subject, which may be marked by a pronoun, a noun, or agreement on the verb. Coding of the third-person subject on the verb in Lele (East Chadic, Afroasiatic, Chad) indicates that the subject of the clause is distinct from the preceding third-person subject (Frajzyngier, 2001). It appears that those differences are due to the default value of the linguistic form, first proposed for the systems of reference by Comrie (1998) extended here to lexical items. It appears that the default referential values of lexical items across languages are not completely accidental, but the issue remains to be explored (see Frajzyngier, 2019).

A study of relative complexity, like a study of typology of functions, requires a non-aprioristic analysis of functions in a given domain. Such an analysis should be based on language-internal data and on relations between the functions in the given language, rather than on some “canonical” definitions of categories.

In what follows I provide sketches of the reference system in a few languages belonging to different families. Each sketch consists of two parts: The first is a description of the structure and functions encoded in the reference system and the second is a description of what functional domains are interacting with the reference system, i.e., functional domains that the speaker has to take into consideration when realizing the functions coded in the reference system. Each sketch is based on first-hand analyses of the language in question. The set of functions within the reference system constitutes the totality of the complexity of the reference system. For analyses that have not yet been published, I provide the argumentation. For other analyses the reader is referred to the appropriate references.

The choice of English as the first language in the description is driven by the fact that I illustrate here the method of presentation and the theoretical approach, and analyses and argumentation are more readily understood when they are based on data familiar to the reader.

REFERENCE SYSTEM IN ENGLISH

The Coding Means in the Reference System of English

Deployment of a noun phrase

Omission of the noun phrase from the environments where it can occur

Subject and object pronouns

Bare nouns in the singular (i.e., nouns without any determiner)

Bare nouns with a plural marker

Articles: definite *the* and indefinite *a*

Demonstratives and determiners: *this, these, that, those*

Possessive pronouns, *my, your*, etc.

Quantifiers: *some, all, any, (a) few*.

Subject and Object

The description of the reference system in English must make a distinction between the functions encoded at the clausal level and the functions encoded at the level of the noun phrase. At the clausal level, all grammatical relations share the function of coding a new participant, marked by a full noun phrase. Another function shared by all grammatical relations is the instruction to identify the participant within previous discourse, within the environment of speech (deixis), or within the listener’s cognitive state. The following example illustrates the introduction of new participants, *animal control* and *fruit trees and bushes*, and the instructions to identify subject and object through the pronouns *they* and *them*, which here happens to be the topic of the message: bears scavenging in the city:

- (1) *Please don’t call **animal control**. **They** are hungry due to climate change. **The fruit trees and bushes** that feed **them** didn’t produce this year.* (website nextdoor.com)

Coreferentiality with the subject of the preceding clause is marked by the absence of the nominal or pronominal subject in the positions in which such subjects can occur.

Inherent Properties of Nouns: Non-entities vs. Entities

A striking characteristic of English is that bare singular nouns occur very rarely in natural discourse. This fact needs to be explained as in many languages there are no constraints on the occurrence of bare singular nouns. The constraint on bare singular nouns explains the syntactic and semantic complexity of the noun phrase in English. Given the importance of this issue for the system of reference cross-linguistically and, more specifically why noun phrases in English appear to be more complex than in other languages, the following discussion includes the state of the art, the hypotheses and the argumentation. I propose that English nouns other than proper names, toponyms, and mass nouns designate “semantic concepts,” similar in properties to the consonantal roots in Semitic languages (Gragg and Hoberman, 2012) and to bare nouns in Mandarin Chinese (Frajzyngier et al., 2020).

State of the Art

The term “bare nouns” in the literature on English (and sometimes in other languages) has in its scope any noun, singular and often plural, that does not have a determiner. Bare nouns in English attracted much attention from generative linguists some 30 years ago because of certain aprioristic assumptions about the structure of the noun phrase that included a determiner as its component (Carlson, 1980; Longobardi, 2001; Delfitto, 2006 and numerous references there). A frequent approach is to describe the function of bare nouns through inferences about their referents in the real or imagined world. Most often mentioned meanings are “kind” or “exemplars of kind” (Carlson, 1980; de Swart and Zwarts, 2009; Le Bruyn et al., 2017).

Payne and Huddleston (2002: 328) propose that NPs such as “*president, deputy leader of the party* [are] bare in the sense that they do not contain a determiner.” The bare role NPs are qualified as NPs by virtue of their being the predicative complements of verbs like *be, become, appoint, elect*. Singular NPs of this kind are

exceptional in that they cannot occur as subjects or objects in a construction where a determiner such as the definite article *the* is required:

I'd like to be president
*I'd like to meet *president/the president*

It would thus appear that the use of the bare noun is determined by the type of predicate.

Quirk and Greenbaum (1973: 73) state that “There are a number of count nouns that take the zero article in abstract or rather specialized use, chiefly in certain idiomatic expressions (with verbs like *be* and *go* and with prepositions).” This is followed by numerous examples of count nouns following the predicates *be in*, *go to*, *travel*, *leave*, *come by*. Interestingly, the table of examples is organized by the types of nouns, which include seasons; some institutions; means of transport; times of the day and night; meals; illness; and parallel structures such as *hand in hand*.

Stvan (2007) reviews the literature concerning the usage of bare singular nouns; concentrates on the use of bare nouns referring to locations, such as *campus*, *cellar*, *sea*, *temple*, etc.; and analyzes their functions through the analysis of various situations referred to by phrases with bare locative nouns.

A Hypothesis Regarding Bare Nouns in English

The present study differs from previous studies in limiting the notion of “bare nouns” to singular, non-mass nouns without determiners and without possessive pronouns. Plural nouns without any determiners belong to an entirely different set, as they code a different function. The present section describes only the function of singular bare nouns, excluding mass nouns. This exclusion is not arbitrary but rather is based on the fact that mass nouns share several syntactic properties not shared with other bare nouns and, more specifically, mass nouns can function as subjects and objects in a large variety of predications.

The reference system in English distinguishes between reference to entities in the real world or in the preceding discourse and nouns that do not refer to entities. Bare nouns in the singular in English represent concepts rather than entities. In order to represent entities, singular nouns in English must have one of the formal means added, such as a plural marker, an article or another determiner, an adjective, a numeral, or a possessive pronoun. Hence the inherent property of bare nouns is a reason for increased formal complexity of the expression and the associated semantic complexity. The evidence that bare nouns represent concepts rather than entities is provided by several constraints on their distribution in English and by semantic outcomes of their deployment.

Evidence From Syntactic Constraints

All discussions of bare nouns in English agree that they cannot serve as subjects or objects in a clause. The question is, why is this so? In many other languages, as illustrated later in this study, there is no such constraint on bare nouns. Therefore, the reason

for the constraint must be a semantic contradiction between the functions of the unit “clause” and a semantic property of bare nouns. In the traditional approach, a clause “is the description of some activity, state, or property.” (Dixon, 2010: 93). Assuming this approach, we can take it that the predicates refer to some states or activities and that noun phrases represent participants in the sense of Lazard (2004). Concepts, as postulated above, are not participants in the event, as they are features in the structure of the vocabulary. This explains the internal contradiction between the function of the clause and the semantic properties of singular bare nouns.

Evidence From Semantic Outcomes of Deployment

One piece of evidence that bare nouns do not represent entities is provided by the contrast between equational predications, where the predicate is a bare noun, and the same types of predications where the predicate is not a bare noun. When the predicate is a bare noun, the outcome of the predication is not another entity but rather the same entity with a new set of properties. When the predication has a non-bare noun as a predicate, the outcome is a new entity. Most examples are from the Corpus of Contemporary American English (COCA).

In clauses of the type *he became professor*, the outcome is just one entity with a new property, that of being professor:

- (2) *Late in 1922, Stern **became professor** of physical chemistry at the University of Hamburg.*
- (3) *In fact, she did not receive one until she was in her mid-fifties, when she was **made professor** and head of the department;. . .*

When the predicate in an equational predication is a determined noun, such a predicate represents a different entity:

- (4) *Felder **is a professor** of pathology and associate director of clinical chemistry.*

There is a set of entities, professors of pathology, and Felder is one of them.

Evidence From Relative Clauses

Relative clauses in English cannot have bare nouns as their heads, regardless of the grammatical relation of the head noun:

- (5) **man who knocked on the door*
**door on which the man knocked*

Any of the utterances in (5) would become grammatical if one were to precede the head noun by a definite article, e.g.,

- (6) *the man who knocked on the door*

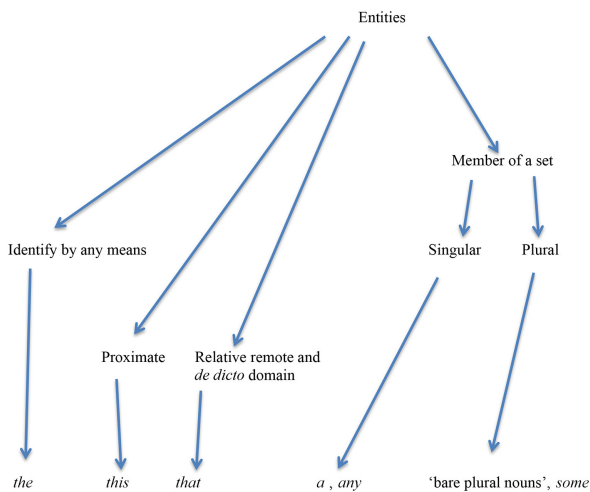
The question is why there is this constraint on English relative clauses. In other languages, e.g., in Polish, the head of the relative clause may be a bare noun:

(7) *Człowiek, który zapukał do drzwi*
 man:NOM REL:M:SG KNOCK:PFV:PST:M:SG to door:GEN
 “The man who knocked on the door”

The hypothesis proposed in this study provides a principled explanation for the constraint in English. Bare nouns in English represent semantic concepts, which are defined in relationship to other concepts in the lexicon. As such, bare nouns belong to the domain *de dicto* in English. The modification of a noun by the relative clause is also a modification in the domain *de dicto*. The coding of the noun already in the domain *de dicto* by a relative clause would constitute a tautology with respect to the function coded.

Identifying the Referent of Entities in English

The following diagram represents a proposed structure for identifying referents of entities in English. The important point of this diagram is that the function of each form is determined not by inferences about the reality that can be obtained from the deployment of one or another form in a particular utterance, but rather from the relationship between various functions within the system. The middle-tier labels refer to functions through which the identity of the reference should be established. The lower tier represents the morphological coding of these functions. The explanation of the functions that need to be explained follows the diagram:



The function “identify by any means” is the same that is traditionally called “definite” in English and is often defined as referring to an identifiable entity (Matthews, 1997: 49). This type of description takes the point of view of a reasonable speaker, who presumably will not use the article *the* if the identity of the referent is not identifiable. But language serves both the reasonable and the unreasonable speakers. The definition proposed here accounts for the function within the semantic system of the language.

Proximate distance: The demonstrative *this* tells the listener to identify the entity, event, state, or even a fragment of speech as

proximate. The demonstrative *that* indicates a relative distance with respect to the speaker when some other entity or situation is more proximate. The important factor here is that the distance is relative rather than absolute. A given absolute distance between the speaker and the referent can be described either by the form *this* or by the form *that*. Distance is therefore not the factor. But if between the speaker and the intended referent there is another referent, even an imaginary one, the intended referent is referred to by the form *that* rather than *this*.

Remote distance: The referent is not in the range of vision of the speaker or listeners but has been mentioned in the immediately preceding discourse:

(8) *I do not like **this** chief because he wants to cut some of my coworkers.*

Relative distance with respect to point of reference: The form *that* city is used because another town has been mentioned earlier:

(9) *My parents used to move all the time; while I was off at college they moved one last time. I lived there my junior year of college, the year between college and grad school, and one summer after grad school. I do not like **that** city at all.*

The form *that* is also used in the *de dicto* domain, i.e., referring to the content of speech, not the speech itself, as explained in Frajzyngier (1991):

(10) *Because once you label yourself a role model, people start judging you, saying you should be this way or that way. And I do not like **that** at all.*

In several unrelated languages, markers that mark the entities more remote from the speaker also mark entities in the domain *de dicto*.

Conclusions About English

The grammatical system of reference in English makes a basic distinction between entities and non-entities. For entities, there are five functions that instruct the listener how to identify the referent: identify the referent by any means (“definite”), proximate referent, relatively remote referent, member of a set. For the functions proximate, remote, and member of a set, the speakers must choose between two numbers: singular and plural. Altogether, the speaker of English must compute seven functions in the coding of reference.

REFERENCE SYSTEM IN MINA

Mina (Central Chadic) is spoken in several villages and settlements in the western part of Northern Cameroon. The main Mina village is Hina-Marbak. The data come from Frajzyngier et al. (2005), but the analyses are new. The reference system of Mina codes some functions that are not coded in the better-known grammatical systems of Indo-European languages, such as deduced reference, switch reference, and remote reference.

The Formal Means of Coding in the System of Reference

Bare nouns;

The determiners *tà* “deduced reference,” *wà* “specific reference,” or *nákàhà* “remote reference,” or *nákàhà* and *wà* at the same time (in that order);

Pronouns;

The omission of nouns or pronouns in the subject or object role;

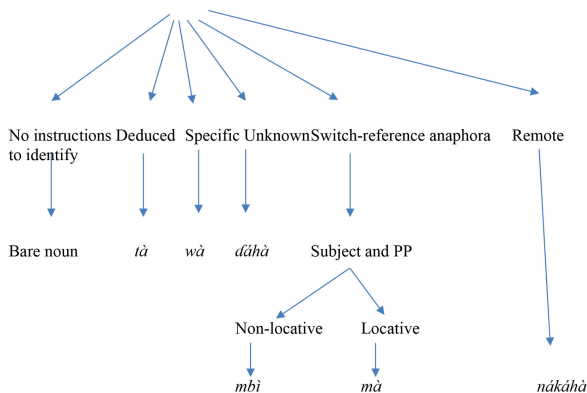
The demonstratives and independent anaphors *m̀bì* “thing,” *m̀à* “there,” *k̀à* “here,” which are distinct from pronouns;

A two-number system in both nouns and pronouns, but no number distinction in deictics or anaphors;

The nouns *hìd* “man” and *m̀bì* “thing,” which code an unspecified human and unspecified non-human entity, respectively. The plural of the noun *hìd* may refer to plural humans as well as to plural animals. The noun *m̀bì* “thing,” but not *hìd* “man,” has been incorporated in the class of pronouns and determiners and undergoes phonological changes that distinguish that class from all other classes of grammatical and lexical items.

Functional Domains in the System of Reference

The following diagram represents the functional domains in the system of reference in Mina. The system of reference in Mina



In what follows I describe each of these functions and their interaction with other functions within the system, arbitrarily starting with the left side of the system depicted above.

No Instruction to Identify the Participant

The fundamental functional distinction in the system of reference in Mina is between (a) no instruction to identify the referent, and (b) instructions as to how to identify the referent. No instruction to identify the referent is coded by the deployment of the bare noun or a noun followed by possessive pronouns. The bare nouns may refer to entities that have never been mentioned before and are unknown to either the speaker or the listener or to entities that have been mentioned in discourse and are well-known to the listener and

the speaker. Here is an example where nouns *t̀àk̀àr* “turtle” and *ɣ̀əm* “water” both occur without a determiner, although they have been mentioned in the preceding discourse. Even the second mention of the noun *k̀ilif-ỳi* “fish-PL” occurs without a determiner:

- (11) *séy t̀àk̀àr t̀il á ǹè ɣ̀əm*
 so turtle leave PRED PREP water
m̀əl m̀əl á m̀əl-á dz̀əb̀əɣ̀
 seize seize 3SG seize-GO five
 “So, the turtle went in the water and caught five [fish].”
séy k̀ilif-ỳi í d̀ám̀d̀ám̀ə í m̀ə ǹj-í
 so fish-PL 3PL good:RED 3PL REL be-STAT
 “So, the fish are good. They are there.”

Deduced Reference

The function labeled “deduced reference” instructs the listener to deduce the referent of a noun from preceding discourse when the referent was not mentioned in the preceding discourse or when there were several potential referents mentioned and the listener needs identify only one for the predication in question. The form *t̀àɣ̀* does not say which particular referent has to be chosen.

The following example ends with the clause *í h̀óyǹə t̀àɣ̀* “they cure it.” The form *t̀àɣ̀*, translated as “it” for lack of a better form in English, could have as its potential antecedent *m̀b̀ig̀iɣ̀* “ceremony,” *m̀əts* “sickness,” or *h̀àɣ̀ák* “village.” Given the context of the utterance, only one of these nouns, *m̀əts* “sickness,” is the antecedent of the form *t̀àɣ̀*:

- (12) *m̀b̀ig̀iɣ̀ ẁàc̀iɣ̀ í d̀ál ǹg̀ám m̀əts*
 m̀b̀igin DEM 3PL do because sickness
k̀ə d̀ál ǹə h̀àɣ̀ák í h̀óyǹə t̀àɣ̀
 INF do PREP village 3PL calm¹ DED
 “This m̀b̀igin [a rite], they do it because there is sickness in the village. They cure it.”

In the following example, there are two groups of potential participants in the event. The determiner *t̀àɣ̀* directs the listener to make a choice:

- (13) *z̀íɣ̀ ǹg̀ul-ỳi p̀ár s̀ulú́d t̀àɣ̀*
 then man-PL other two DED
í ǹd-áhà b̀áhá
 3PL go-GO again
ǹd-á m̀áb̀àr m̀b̀ír b̀áhá k̀ə m̀əl t̀àɣ̀
 go-GO lion leap again INF seize DED
 “Later, when the two other men arrived, the lion jumped to catch them.”

If the preceding discourse contains only one noun phrase, the form *t̀àɣ̀* does not refer to that noun phrase but to something else related to that noun phrase.

In the next example, the Koran is the object of the first clause. However, the only overt object marker in the second sentence is the marker *t̀àɣ̀*. Because the reduplication of the verb *ǹáz* “throw” indicates a repeated action, the antecedent of *t̀àɣ̀* cannot be the Koran itself, which is one entity, but must be some plural object

¹This item is borrowed from Fula.

associated with the Koran. This object can only be the pages of the Koran, even though the pages themselves have not been overtly mentioned. The use of *tàŋ* thus instructs the listener to deduce the referent for the object:

- (14) *ḡət á ḡət déftə ngən*
 take 3SG take Koran (F.) 3SG
 “He took his Koran.”
pàts ntá nàz nàz nàz nàz á nàz
 took one throw throw throw throw 3SG throw
tàŋ á nè yəm wàhíŋ
 DED PRED PREP water DEM
 “He took one [page] after another and threw them upon the water.”

The Category “Specific”

The domain for which the general term “specific” is given here instructs the listener to identify the referent (a) within the environment of speech, including the immediately preceding discourse, (b) as the topic of the discourse, and (c) within the preceding proximate discourse. The category “specific” is coded by the marker *wà*, whose phrase-final forms are *wàcín* or *wàhín*:

- (15) *nòk kè dál žì vàŋgáy kè kázŋ làkwát wàcín*
 1PL INF do then how INF cross river DEM
 “How are we going to cross this river?”

The Category Unknown

The category “unknown” tells the listener that the speaker does not want the listener to search for the identity of the referent. The category is coded by the verb of existence *dáhá* (phrase-final form) or *dá* (phrase-internal form). Consider the following example:

- (16) *kè nàz ngùl á bíŋ dáhá*
 INF leave man PRED house exist
 “She abandoned a man in the house.”

The evidence that the form *dáhá* codes an unknown entity is provided by the fact that if the entity is known, the form *dáhá* cannot be used. Thus, if one adds the third-person possessive pronoun *ngəŋ* after the noun *ngùl* “man” in the above example, one cannot use *dáhá*. The reason for the ungrammaticality of (18) is quite simple: If the man is the woman’s husband, the house where they are is also his and her house, and therefore it cannot be an unknown house:

- (17) *kè nàz ngùl ngəŋ á bíŋ *dáhá*
 INF leave husband 3SG PRED house *exist
 “She abandoned her husband in the house.”

The Switch-Reference Anaphora

Mina has an elaborate system of coding switch-reference anaphora. The term switch-reference anaphora in this study refers to the function of coding the referent of a participant or place as one that has been mentioned before but not the one that was mentioned in the immediately preceding clause. The term switch reference, as used here, therefore applies to a broader range of relations than does the usual understanding of switch reference to the subject, as known

in North American Indian and New Guinea languages. The switch-reference anaphora has three subfunctions: one for the subject and complement of a preposition, with the further division into inherently locative and inherently non-locative nouns, and one for the complement of the preposition for inherently locative nouns. For inherently non-locative nouns, the switch-reference anaphora function for the subject and object of the preposition is coded by the form *mbí* (*mbə* phrase-internal, *mbéŋ* phrase-final), glossed as S.R.ANAPH for “switch-reference anaphora.” The form always functions as the head of the noun phrase, i.e., it is never a determiner. The antecedent of the switch reference marker may be a noun phrase or a state or an event described by a proposition or by a larger chunk of discourse.

The switch-reference marker for inherently locative nouns is *mà* (underlying form) and *mècín* or *mèhín* (phrase-final form). The following examples illustrate the deployment of switch-reference marker in the function of the subject and complement of preposition. There are no examples of coding the object with the anaphor *mbí*. In the following examples, switch reference has as its antecedent somebody who was mentioned in the previous discourse:

- (18) *báy wílè á dàmù mbí*
 chief still PRED bush S.R.ANAPH
nd-á ḡət wədə
 go-GO take food
 “The chief_i is still in the bush. He_j came to take the food.”

- (19) *hidi mindéŋ à n ká bə̀t də̀və̀r*
 man other 3SG PRED INF make hoe
gə̀ gə̀ rə̀ sùlú́d ábə̀ mbéŋ
 ten ten hand two ASSC 3SG
 “Another person will make twenty hoes with that.” (*gə̀* comes from *gə̀b* “ten”).

- (20) *séy kə̀-yíí dí zə̀ ngə̀ŋ ká*
 then cow-PL put EE 3SG POS
á nè mbéŋ
 PRED PREP 3SG
 “Then the cows, he kept them, for himself.”

Since the locative anaphor *mà* is used with inherently locative nouns, it modifies nouns without the locative preposition.² In the following example, the noun *bíŋ* “room” is mentioned in the first clause. The subsequent mention of the room in the third clause is marked by the form *mà* [both instantiations are bolded in examples (21) and (23)]:

- (21) *tíl á ndə̀ zə̀ bíŋ*
 depart 3SG go EE room
à n mì bíŋ dzáŋ á dzáŋ ká
 3SG PREP mouth room close 3SG close POS
 “He went to the room and closed the door.”

²Inherently locative nouns are not marked by a locative preposition in locative predications in Mina (Frajzyngier et al., 2005).

- (22) *báhámàn lù á lùw-á-ŋ nɔ̄ kʒámáy*
 Bahaman say 3SG say-GO-3SG PREP stick
nákà wà
 REM DEM
 “Bahaman spoke to the stick.”
- (23) *kʒámáy wà mɔ̄l á mɔ̄l-á-ŋ*
 stick DEM catch 3SG catch-GO-3SG
ndɔ̄ ngɔ̄n bíŋ mácíŋ
 beat 3SG room ANAPH
 “The stick started to beat him in the room.”

Remote Identification

Remote anaphora is marked by the form *nákáhà*, which follows the noun. The form *nákáhà* can modify the object or can function as a complement of a preposition. The form can be used even if it did not have antecedent, referring to some point in time or the event that the listener should use as a referent:

- (24) *séy tíl ndɔ̄ dzáŋ í dzáŋ kilíf*
 so go go find 3PL find fish
gwád ángɔ̄ nákáhà
 plenty like REM
 “So, they went and found a lot of fish, as previously.”
- (25) *séy dɛw tɔ̄tɔ̄ kɔ̄ mɔ̄na nákà mɔ̄kʒɛkʒɛ*
 so sit 3PL like DEM REM before
 “They remained as before.”

Unlike other determiners, the remote reference marker may be followed by the deictic *wà* coding the reference as known:

- (26) *nd-á zɔ̄m zɔ̄m nákà wà zá*
 go eat eat REM DEM EE
 “They returned and ate that one” (i.e., the guinea fowl mentioned five sentences earlier).
- (27) *fúu tàŋ hidì gɔ̄nák díyà bɔ̄ŋ séy í hán*
 all DED man black put think so 3PL cry
rá mbɔ̄ nákà gárkʒàw wácíŋ séy díyà
 D.HAB child REM disobedient DEM so start
rɔ̄ jíb í jíb hós á útɔ̄ wál
 dig hole PREP hole arrive PRED house woman
nákà wácíŋ mɔ̄ bɔ̄t wɔ̄zì nákà wácíŋ
 REM DEM REL take children REM DEM
 “All the people started thinking. Then, they were crying. The disobedient child started digging a tunnel to the house of the woman who took those children.” (*hidì gɔ̄nák* “man black” = “man,” *jíb í jíb* “tunnel (hole in a hole).”

The term “remote” is a relative term, indicating that between the potential antecedent and its repetition may be several other nouns whose referents may have the same role. In the following example, the noun *ngɛf* “feather” is followed by other nouns, such as *bàkátàr* “bag,” *kúhú* “fire,” *ndri* “corn,” and the subject, *gàmták* “chicken”:

- (28) *séy gàmták báhà wérèh wérèh séy*
 so chicken again clever clever so

- bɔ̄t ngɛf ngɔ̄n tú gúráy tú gúráy bɔ̄k*
 take feather 3SG GEN large GEN large put
á nɔ̄ kúhú séy tíl ngɔ̄n nɔ̄
 PRED PREP fire so enter 3SG PREP
bàkátàr díy-á zɔ̄m ndri díy-á bám
 bag put-GO eat corn put-GO eat
kʒì tó n bákátàr tùwɔ̄d kà
 meat GEN PREP bag finish POS
 “So the clever chicken took his large feather, put it into the fire. He himself entered into the bag, started to eat sorghum, started to eat meat [and] finished everything that was in the bag.”

When reference is made to the noun *ngɛf* “feather” in the next sentence, *ngɛf* is followed by *nákáhà* because there were several noun phrases between its previous and the current mention:

- (29) *kwáyàŋ tí syì ngɛf nákáhà wɛcíŋ díy-à*
 squirrel see COM feather REM DEM put
njif á njif grá kʒì tó gàmták
 smell 3SG smell like meat GEN chicken
mɔ̄ mɔ̄sáw-yí zà zidép
 REL grill-STAT EE already
 “The squirrel saw that those feathers smelled like the flesh of the grilled chicken.”

Summary for Mina

As illustrated in the diagram above, Mina has seven functions through which the speaker may direct the listener to identify, or not identify, the referent of a noun in discourse. These functions include information about how the listener should go about identifying the referent; about the role of the referent in the proposition, whether subject or not; and about the semantic property of the referent, whether locative or not. The categories human or non-human, gender, and number, which serve as functions and coding means in many other languages, do not play a role in the reference system of Mina.

REFERENCE SYSTEM IN POLISH

The Formal Means of Coding

The formal means of coding within the reference system of Polish include:

Nouns

Bare nouns in Polish have different inherent properties than bare nouns in English or Mandarin. Bare nouns in Polish always represent entities rather than concepts. The relation of this property of nouns to the overt inflectional marking of gender remains to be thoroughly examined. As a result of this property of nouns, Polish does not have any markers whose function is to convert concepts into entities. On the other hand, it has periphrastic means of converting entities into concepts.

Numeral *jeden*: “one” in non-literary Polish and a corresponding adjective *pewien* “certain” in literary Polish.

Gender and number are coded on verbs (gender and number of the subject), adjectives, numerals, demonstratives, and determiners. Polish has a three-gender system in the singular and a two-gender system in the plural. The gender system in the plural

does not correspond to the gender system in the singular, hence one could talk about a five-gender system marked by a variety of morphological means. In addition, within the class of masculine nouns in the singular there is a distinction between animate and inanimate masculine, and personal vs. non personal in the plural, in effect adding two more genders. The number distinction is binary, with the singular unmarked and the plural marked.

The coding of gender and number on verbs and nominal categories, often referred to as “agreement,” is an independent coding means within the system of reference rather than a mechanical outcome of the presence in the clause of some “trigger” noun having the features of gender and number (Frajzyngier and Shay, 2003). For a different approach see Roberts (2019) and Corbett (2006).

Person

There are three persons in the singular coded in the pronominal system and on the verb. In the plural there are two persons, human masculine and all others (i.e., nouns that in the singular are masculine non-human, feminine, and neuter). The verb codes a two-gender distinction in the second person and a three-gender distinction in the third person. The pronominal system does not code the gender distinction in the second person. Verbs do not code gender distinction in the present tense and one type of future. The coding of the gender, number, and person of the subject on the verb is obligatory regardless of whether there is a nominal or pronominal subject in the clause. Moreover, the coding of the subject on the verb involves more distinctions of gender in the second person in the past tense than could be represented by the pronouns.

Question Words With the Suffix -ś

The term “question words” refers to morphemes referred to in Polish grammars as interrogative pronouns, such as *kto* “who,” *co* “what,” *który* “which one (M.),” *jaki* “what kind (m.)” (with corresponding forms in feminine and neuter, and five plural paradigms). From all of these words one can derive a noun through the addition of the suffix *-ś*. The nouns so derived indicate a participant having the defining feature of the question word, such as human, non-human, or attribute, but otherwise unknown to the speaker.

Pronouns (for a taxonomy of pronouns see Laskowski, 1984).

In the first- and second-person singular and plural there is no distinction of gender (all forms are represented here in standard orthography): *ja* (1SG), *my* (1PL), *ty* (2SG), *wy* (2PL). The third-person pronouns are distinguished for gender: *on* (M:SG), *ona* (F:SG), *ono* (N:SG), *oni* (M:HUMAN:PL.) and *one* (PL for all other referents).

Proximate deictic and an unrestricted determiner series, *ten* (M), *ta* (F) *to* (N), *ci* (HUMAN MASCULINE, PL), and *te* (the rest of the nouns).

Demonstratives and determiners of the series *tam-ten*, *tam-ta*, *tam-to* (morphemic division inserted, but not marked in Polish orthography).

Anaphors only, (possibly limited to the literary variety): *ów* (M), *owa* (F), *owo* (N), *owi* (PL:M:HUMAN), and *owe* (all others). This series does not have a distinction between proximate and remote mention.

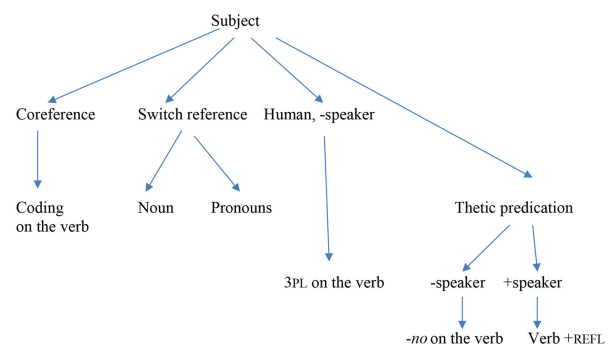
All demonstratives and anaphors, as well as nouns derived from question words, can function on their own as arguments or adjuncts in the clause. All demonstratives, anaphors, and nouns derived from question words can also function as determiners of nouns.

Case marking of nouns, pronouns, etc., is not only a means to code the semantic or grammatical relationship between the predicate and noun phrases or relationships between noun phrases, but also has an important function in the coding of reference. The anaphoric or cataphoric function associates (“binds”) the marker in a given clause with a noun having the same case in the preceding or the following discourse. It is case marking that enables a variety of markers to function as a coding means within the system of reference.

The Overall System of Functions

Within the reference system of Polish one needs to make a distinction between (a) reference to subject and (b) reference to all other grammatical and semantic relations between the verb and noun phrases. In particular, for the subject there is a tripartite division between coreference with the immediately preceding subject, switch reference with respect to the immediately preceding subject coded by the deployment of noun or subject pronouns, and the coding of unspecified human subject. For this last category there is a further distinction between the forms that exclude the speaker and the forms that allow the inclusion of the speaker. The coding on the verb is an independent coding means rather than an agreement system, as evidenced by the fact that it codes more functional distinctions than are coded on pronouns. Thus, the distinction of masculine and feminine gender in the first- and second-person singular and plural is not coded on pronouns. It is, however, coded on the verb in the past tense.

Reference to the subject in Polish:



What follows is an explanation of each of these functions. The term “-speaker” indicates exclusion of the speaker and the term “+speaker” indicates potential inclusion of the speaker.

Thetic Predication—Excluding the Speaker

Thetic predication in Polish indicates the event only from the point of view of what happened, not from the point of view of an agent or an experiencer. Such predication is coded by the verb with the suffix *-no* in the past tense. In the present tense, thetic predication is coded by the verb in third person along with

the reflexive marker. One cannot add nominal or pronominal subjects to such predications.

The following example has the suffix *-no* on the verb *rozwi-
no* “they were setting up”:

- (30) *Właśnie rozwija-no namiot cyrkowy.*
just set.up-TP tent:ACC circus:ADJ
“They were just setting up the circus tent” (Sławomir
Mrożek, *Tygodnik Powszechny*, 22/1983 via [https://
lekturygimnazjum.pl/artysta-slawomir-mrozek-tekst/](https://lekturygimnazjum.pl/artysta-slawomir-mrozek-tekst/)).

Thethetic predication implies human participants only. The following clause, which describes eating habits at a zoo, can have only human consumers in its scope, not the animals that also eat in the zoo.

- (31) *W zoo jada-no tylko świeże owoce i jarzyny*
in zoo eat-TP only fresh fruits and vegetables
“In the zoo only fresh fruits and vegetables were eaten.”

Thetic Predication Including the Speaker

The possible inclusion of the speaker in thethetic predication is coded by the third-person singular neuter form of the verb followed by the reflexive marker *się*. In the past tense, the form of the verb has the suffix *-ło*. No nominal or pronominal subject can be added to such clauses:

- (32) *Zakładalo się jedną parę skarpetek więcej*
put.on:PST:N REFL one pair socks more
i to
CONJ DEM
rozwiązywało problem, bo za rok już
solve:N:PST problem because in year already
pasowały jak ulał.
fit:PST:PL:N perfectly
“One would put on one more pair of socks, and that used to solve the problem, because in a year, they [shoes] would fit perfectly” (<https://podlaskisenior.pl/jak-sie-dawniej-ubierano/>)

New Subject

A new subject in discourse, which in one way or another will be referred to in the subsequent discourse, is coded through the overt coding of the noun phrase. Such a noun phrase can consist only of a bare noun, as illustrated in the next section.

Coreference

Verbs in the past, present, and future tenses obligatorily code the person, number, and in some tenses gender of the subject, regardless of whether the clause has or does not have a nominal or pronominal subject. In each case, the coding of the subject on the verb indicates coreference with the immediately preceding subject. In the following example, the new subject *dyrektor* “director” is followed by the verb *przyjął* “received.” The last clause, *gdzie urzędował* “where he worked,” does not have a nominal or pronominal subject. It codes coreference with the subject of the preceding clause through the coding on the verb:

- (33) *Dyrektor przyjął go na świeżym*
director:NOM receive:PFV:PST:3SG:M 3SG:M on fresh
powietrzu, gdzie urzędował.
air where work:IPFV:PST:3SG:M
“The director₁ received him outdoors, where he₁ worked”
(Sławomir Mrożek, *Tygodnik Powszechny*, 22/1983 via
<https://lekturygimnazjum.pl/artysta-slawomir-mrozek-tekst/>)

Unspecified Human Subject That Does Not Include the Speaker

Unspecified human subject is often coded by the third-person plural masculine subject on the verb, without any pronouns (a necessary condition). Here is an example of the first line of a narrative, hence there are no potential nominal antecedents. The relevant verb is *zaprosyli* “they invited”:

- (34) *Znowuś jednego młynarza zaprosyli na*
one.time one:GEN miller:ACC invite:PFV:3PL:M on
kśćiny na drugom wieś
christening on another village
“One day, they invited a miller to a christening in another village . . .” (A better translation could perhaps be “A miller was invited to a christening in another village.”)
[Nitsch, 1960: 144. This and other dialectal examples are transcribed as in Nitsch (1960)].

In contemporary literary Polish, the third-person plural human masculine can also code an unspecified human subject. The referent of such a subject could be masculine or feminine:

- (35) *W powszednie dni wszyscy ubierali się*
in ordinary days everybody dress:PST:3PL:M REFL
skromnie, nawet biednie.
modestly even poorly
“On ordinary days, everybody dressed up modestly, even poorly” (<https://podlaskisenior.pl/jak-sie-dawniej-ubierano/>)

In the present tense, the third-person plural coding on the verb, again without any pronouns (a necessary condition), codes the unspecified human subject in both literary and non-literary Polish. Here is an example from non-literary Polish. The utterance is the first line in the narrative, hence there are no potential antecedents. The relevant verb in the following example is *godajom* “they say”:

- (36) *Godajom, że tero to każdy gospodarz*
say:PRES:3PL COMP now COM every farmer
bogoć
rich.man
“They say that nowadays every farmer is a rich man”
(Nitsch, 1960: 188, recorded in 1920).

Contemporary literary Polish will also have the verb in the third-person plural present tense:

- (37) *Powiadają, że teraz to każdy gospodarz*
 say:PRES:3PL COMP now COM every farmer
jest bogaczem
 be:PRES:3SG rich.man:INS
 “They say that nowadays every farmer is a rich man.”
- (38) *Gdzie trzciny, tam woda, powiadają.*
 where reed there water say:PRES:3PL
 “Where there are reeds there is water, they say.” (NKJP).

Switch Reference to Previously Mentioned Subjects

As Polish obligatorily codes the person, gender, and number on the verb, subject pronouns are deployed to code switch reference to the subject that was mentioned previously in discourse or that is imagined to have existed in the preceding discourse, or focus on the subject that has been previously mentioned (Frajzyngier, 1997):

- (39) *Wystrzelisz, on upadnie.*
 shoot:2SG:FUT 3M:SG fall.down:3SG:FUT
 “You will shoot, and he will fall down” (Jarosław Iwaszkiewicz, *Brzezina*, via NKJP).
- (40) *Mnie też się wydawało, że ładna to ona nie jest.*
 1SG:ACC also REFL appear:N:PST COMP pretty DEM
 3F NEG be
 “I also had an impression that pretty she is not” (NKJP).

Summary for the Reference on the Subject

The coding of subject in Polish is driven by five functions, each of them coding a different class of entities within which the subject is to be identified, and by one function that does not include the subject. Nouns that code the subject, third-person pronouns, and coding on the verb distinguish between two numbers, singular and plural. Three genders (animate and inanimate masculine, feminine and neuter) are coded in the singular, while two genders (human masculine vs. all others) are coded in the plural. Since genders in the plural distinguish different functions from the genders in the singular, one needs to postulate the existence of five (with one sub-gender in the masculine) rather than three genders in Polish. The coding of the subject alone in Polish includes five functions with respect to type of reference, and for each function the speaker must make a choice between two numbers and five genders. Including thethetic predication for the subject alone, the speaker has to make a choice between thirteen possibilities.

The following discussion describes the functions that apply to any grammatical role within the clause.

Do Not Identify the Referent

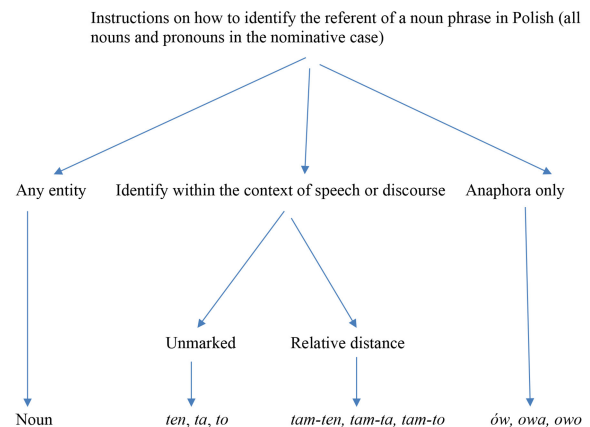
Polish, in both literary and non-literary variety, has a means to inform the listener that the identification of the referent is irrelevant for the following discourse. In contemporary literary Polish this function is coded by the form *pewien* “certain” and, more rarely, *jeden* “one,” with its

masculine, feminine, and neuter forms all declined for number and case:

- (41) *Wu jednygo gospodoza swuzyw Mac'ek*
 at one:GEN farmer served Mac'ek
 “One farmer had a helper named Maciek.” (Nitsch, 1960: 240).
- (42) *Była jedna baba bardzo stara, już*
 be:SG:F one:F old woman very old already
 pewnie do sta lat miała
 perhaps to hundred years have:PST:SG:F
 “There was once a woman, very old, possibly 100 years old” (Nitsch, 1960: 289).

Identifying the Referent of a Participant

The referent of any participant in a proposition or in any grammatical relation can be identified through the following functions:



The demonstratives of the series *ten* and the anaphors of the series *ów* can occur alone or can function as determiners. Both nouns and pronouns must be marked for their grammatical relation with the verb. The four reference functions are facilitated by the existence of five genders and six case markers, which increase the number of forms but provide a more fine-grained identification of the referent.

An Entity

An entity in Polish is coded by a singular or plural form of the noun without any determiners. Topolińska (1984) describes a large number of potential inferences (not calling them as such) that one can draw from the use of bare nouns in Polish. The important fact about bare nouns in Polish, unlike bare nouns in English or Mandarin Chinese, is that they do not code concepts unless a concept, e.g., as derived from the verb, is their referent.

The evidence of the entity function of bare nouns in Polish is provided by the fact that they behave as arguments and adjuncts, in exactly the same way as proper names of people and toponyms, i.e., nouns that by their inherent properties represent unique entities in

any given situation. Here is an example: In the following fragment, the nouns *pies* “dog,” *obserwacja* “observation,” *kobieta* “woman,” *właściciel* “owner,” *człowiek* “man,” and *książeczka* “booklet,” are all mentioned several times, each time without any determiner, in the same way as the toponym *Legionów*:

(43) *W Legionowie prowadzona jest obserwacja psa,*
in Legionów conduct:PASS be:PRES observation
dog:GEN

który 1 marca ugryzł
REL:SG:M:NOM 1.March:GEN bite:3:PFV:PST:SG:M

kobietę na ul. Reymonta.
woman:ACC on str. Reymont:GEN

“In Legionów, they have under watch a dog that on March 1 bit a woman on Reymont Street.”

...

Właściciel psa, który pogryzł
owner:NOM dog:GEN REL bite:PL:PST:PFV:M:SG

człowieka, musi pokazać
man:ACC must show

aktualną książeczkę szczepień.
valid:ACC booklet:ACC vaccination:PL:GEN

“The owner of the dog that bit a person must show a valid book of vaccinations.”

Potem pies musi przejść kwarantannę, tzn.
afterwards dog:NOM must undergo quarantine:ACC i.e.

odbyć trzy
make three

wizyty u weterynarza.
visits at veterinarian:GEN

“Afterwards, the dog must pass quarantine, i.e., must make three visits to a vet.” (<http://nkjp.pl/poliqarp/nkjp1800/query/4/>)

Identify the Referent Within the Context of Speech or Discourse

The instruction to identify the referent through previous mention is coded by two series of demonstratives: *ten*, the unmarked function, and *tam-ten*, the marked function. The function “identify within the context of speech or discourse” encompasses identification through deixis, anaphora and cataphora, deduced reference, and a host of other situations. It does not tell the listener which specific context to choose for the identification of referent. The context is always within the range of knowledge of the speaker and the speaker’s presupposition about the range of knowledge of the listener:

Deixis

(44) *wiecie co to jest ten sweterek*
know:2:PL what DEM:N be:3SG DEM:M sweater:DIMIN
“Do you know what it is, this sweater?” (NKJP).

Inference from the previous mention:

(45) *Nie możemy przyjąć zwrotu, jak nie masz paragonu.*
NEG can:PRES:1PL accept return if NEG
have:2SG:PRES paragon
“We cannot accept the return(s) if you do not have a paragon [sales slip]” (NKJP).

A co to jest ten pa...

(46) CONJ:a what DEM be:PRES:SG DEM:M pa...
“And what is it this pa...” (NKJP) (the speaker did not complete the word “paragon”).

The referent of the form *ten* may be deduced from the previous discourse. In the following example, the speaker talks about an event during the First World War. He situates the pre-battle positions of various armies and uses a demonstrative of the series *ten* before the noun *voda* “water, river.” Obviously, the water in question is not in the environment of speech (the recording was made many years after the war). It has to be deduced from the deployment of the form *tam* “there,” which just indicates a place other than the place of speech:

(47) *Tam stały Prusy nad tom wodom.*
there stand:3PL:F Prussians on DEM:F:INS water:INS
“Over there, the Prussians were standing, near this water.”

Voda s’e nazywała Bzura
water REFL name:3F:SG Bzura

“[The] water was called Bzura” (Note that the second mention of the noun *voda* “water” has no determiner.) (Nitsch, 1960: 269).

The function of identifying the referent within the context of speech or discourse relative to the place of speech or relative to the last mention is coded by the form *tamten*, which, like all other markers listed, can be the sole member of the noun phrase or a determiner. The relevant forms are glossed as R.DEM for “relative demonstrative.” The crucial element in the function of the demonstrative *tamten* is that it is relative with respect to some other referent and that it is not an absolute indicator of the distance. In the following examples, the two sides are defined relative to the wall that separates them, as seen from the point of view of the speaker:

(48) *Panie, widzisz ten mur? Tu jest ta strona.*
Sir see:2SG:PRES DEM:M wall here be:3SG DEM:F
side

“Sir, do you see that wall? Here is this side.”

A tam jest tamta strona.

CONJ there is R:DEM:F side
“And there is that side” (NKJP).

(49) *Wiedziałem i to już mi na cale życie zostało.*
know:1SG:PST CONJ DEM:N already 1SG:DAT for
whole life remain

“I knew, and that remained forever in my life”.

tamta strona nie jest straszna.
 R:DEM:F side NEG be frightening
 “The other side is not frightening”.

Anaphora Only

The demonstratives and determiners of the series *ów* (M:SG), *owa* (F:SG), *owo* (N:SG), *owi* (HUMAN:M:PL), and *owe* (plural determiner for remaining nouns) indicate that the referent has to be identified from the previous discourse, either as previously mentioned or deduced from the previous discourse. A cursory look at the collection of Polish dialectal texts did not result in a single instance of the use of any of these markers. It appears, but again data are not easily accessible, that the marker occurs only in the written medium of the literary varieties of Polish. The antecedent is bolded and underlined, and the determiner phrase is bolded:

- (50) *Panie Staszku, mąż mówił, że*
 Mister Staszek husband say:IPFV:PS:3SG:M COMP
pokazuje się Pan
 show:3SG:PRES REFL Sir
*w towarzystwie **pięknej damy.***
 in company:GEN beautiful lady
 “Mr. Staszek, my husband tells me that you can be seen in the company of a beautiful lady.”
No to pewnie też powiedział, kim
 well COM probably also say:PFV:PST:3SG:M who:INS
*jest **owa dama.***
 be:3SG DEM:F lady
 “Well, so he probably also told you who that lady is.” (Jan Grzegorzczak, *Chaszcze*, via NKJP).
- (51) *Chwalcy kapitalizmu, **owi** tak dobrze opłacani*
 glorifiers capitalism:GEN DEM:PL so well paid
przez państwo
 by state
naukowcy i politycy, nigdy nie byli
 scientists CONJ politicians never NEG be:PL:M
tutaj klientami.
 here clients:INS
 “The glorifiers of capitalism, those well-paid scientists and politicians, were never clients here.” (Bronisław Świdorski, *Asystent śmierci. Powieść o karykaturach Mahometa, o miłości i nienawiści w Europie* via NKJP).

Conclusions About Polish

Together with the function of coding reference of the subject, a speaker of Polish has to take into consideration nine functions in the domain of reference. For each function the speaker also has to consider the fact that each marker may have variants of five genders and, in reference to relations other than the subject, five grammatical and semantic functions marked by case.

REFERENCE SYSTEM IN MANDARIN CHINESE

The discussion of the system of reference in Mandarin just summarizes the hypotheses and argumentations proposed in Frajzyngier et al. (2020).

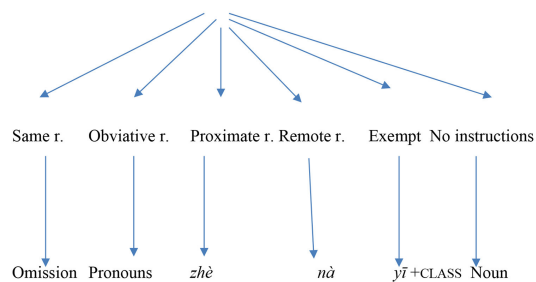
The Formal Means of Coding of Reference in Mandarin:

- Bare nouns
- Proper names and toponyms
- Pronouns
- Omission of nouns or pronouns from the environments where they may be inserted
- Demonstratives *zhè* “proximate this” and *nà* “remote that”
- Classifiers occurring with numerals alone (glossed as CLASS)
- Nouns modified by demonstratives, numerals, classifiers and the marker *yī* “one” + CLASS.

Functions Through Which the Participant Is Identified

Instructions on how to identify the participant in a proposition (“r.” is short for “reference”):

Instructions on how to identify the participant in a proposition (‘r.’ is short for ‘reference’):



The function labeled “same reference” instructs the listener to identify the referent as one of the following: (1) a referent belonging to the speech situation, which could be the speaker, the listener, or even a third person; or (2) a referent that may have been mentioned in the immediately preceding discourse. This function is coded by the absence of a noun or a pronoun in the syntactic slot in which a noun or pronoun might occur. This coding means is labeled as “omission” in the above diagram.

The function labeled “obviative reference” tells the listener that the referent is different from the one that was mentioned most recently but has nevertheless been mentioned in the preceding discourse. The “obviative reference” function is coded by the deployment of pronouns.

The function “proximate reference in space and time” has two subdomains: (1) reference to an entity present in the environment of speech, and (2) reference to an entity that has been previously mentioned but mentioned by a different noun. This function is coded by the proximate demonstrative *zhè* “this.”

The function “remote reference in space and time” also has two subdomains: (1) remote deixis in time and space, and (2) reference to an entity or a proposition mentioned before another entity was mentioned. This function is coded by the remote demonstrative *nà* “that.”

Exemption of the noun from further identification: This function is marked by the numeral *yī* plus the classifier that

is appropriate for the referent. This function is, in a way, a counterpart to the use of bare nouns, which leave the interpretation of the identity to the listener.

The function labeled “No instruction” does not provide the speaker with information on how to identify the referent. This function is marked by the deployment of a noun. This function leaves the identification of the referent up to the listener, involving the use of the bare noun. Bare nouns do not tell the listener how to identify the referent. The evidence for the hypothesis about the function of this coding means is provided not by the analysis of individual instantiations of bare nouns in some clauses but rather by the fact that a variety of grammatical markers can be added to bare nouns to constrain the listener’s interpretation. Some bare nouns, such as proper names and toponyms, have unique referents, while other bare nouns have a large set of potential referents.

Conclusions About Mandarin

Mandarin Chinese codes six functions within the system of reference. These functions only partially overlap with the functions composing reference systems in other languages.

REFERENCE SYSTEM IN A SINO-RUSSIAN IDIOLECT

Basic Information on Sino-Russian Idiolects

The Sino-Russian idiolects are formed by individual Chinese immigrants to the Far East of Russia for communication with Russians. These idiolects are not used for communication within the family or with other Chinese immigrants. The term “Sino-Russian idiolects” is specifically restricted to languages of immigrants who did not have any formal instruction in Russian, or at most very minimal instruction. Each speaker in effect forms her or his own system. The present description is based on Frajzyngier et al. (2021).

The lexical items in the Sino-Russian idiolects may distinguish between verbs and non-verbs, but often there is no categorial distinction between lexical items. All lexical items and the coding means that have segmental realization are borrowed from Russian with no functional distinction of inflectional marking. No Sino-Russian idiolect has an inflectional system on verbs or nouns and there is no gender or number distinction. The only grammatical coding means are intonation, pauses, pronouns, one demonstrative, prepositions, and a few particles. There is no distinction between subject and object, nor is there a coding of semantic relations other than those that are not expected from the semantic properties of the verb. Those semantic relations are coded by prepositions. A common typological feature of various idiolects is the antecedent-comment relation (not to be identified with the topic-comment relation). The predicate, whether verbal or non-verbal, often occurs in clause-final position. In clauses with two participants, the more agentive precedes the less agentive.

The formal means in the coding of reference are:
The deployment of a noun (phrase),

Pronouns

The omission of a noun phrase or a pronoun,
The deployment of the demonstrative *’eta* “this” (with a variety of phonetic realizations, including *’eda*), either alone or as a determiner of a noun.

Functions in the System of Reference of Sino-Russian Idiolects

The functions through which the listener is expected to identify the referent of the noun phrase are: new participant; previously mentioned participant in the same role in the immediately preceding clause; switch reference; deixis; and unknown entity. The locative adverbs *z’es* “here” and *tam* “there” code reference to the place of speech, as broadly understood, and the place other than the place of speech. In what follows is a brief description of three functions. For a full description with a considerably larger number of examples see Frajzyngier et al. (2021).

New Participant in Discourse

New participants in discourse are marked by bare lexical items whose referent could be an entity, corresponding to nouns, or a property concept (/ indicates shorter pause, and // indicates longer pause):

Boris (the speaker’s pseudonym stands for the idiolect from which the example was taken):

- (52) *v’ixaməj n’i vxaməj n’i ab’i’zaat’it*
day.off NEG day.off NEG obligatory
“[The difference between] the day off and not the day off is not obligatory.”

The term “omission” refers to the omission of a constituent from a clause in which the constituent can occur. The omission of a noun or pronoun leaves the interpretation of the omitted entities to the listener’s interpretation. That interpretation is in turn based on the ongoing discourse, on the environment of discourse, and on other constituents included in the utterance.

The fundamental principle in the system of reference in several idiolects is that if a participant and its semantic role—the two necessary components of this condition—can be deduced from the previous discourse, from the environment of discourse, or from constituents of the clause, such a participant is not overtly coded by any means. From this principle it follows that whenever a noun phrase is included, it represents a new participant. Here is an example: In the first utterance a nominal participant, *mu’st’fina* “man,” is mentioned for the first time. In the second utterance there is no nominal or pronominal argument, although the participant is the same as in the first utterance. The second utterance does not have a predicate either. In the third utterance, another participant is introduced, namely *’matf’ik* “boy”:

Slava

- (53) *vət/ mu’st’fina sabi’rajə//*
PRS man:NOM gather:3SG:PRES³
“Here, a man is picking up [pears].”

³Glosses represent Russian, not Sino-Russian, inflectional marking. Although the marking is not productive in Sino-Russian idiolects it is included for future investigation of alternative hypotheses.

- (54) *tr'i ɪs'etkaj u'ʒɛ//*
 three string-bag:INS already
 “[He has picked up] three baskets [of pears] already.”
- (55) *əəə// i'd'ot/ ma 'matʃik//*
 eh go:3SG:PRES boy[ERR] boy:NOM
 “A boy is walking.”

- i uʒbal//*
 CONJ:i fall:3SG:M:PST
 “At this moment he turned his head around, looked and fell down.”

Switch Reference Within Discourse: The Function of Pronouns

The function of pronouns in Sino-Russian idiolects is to code a change of topic/subject in comparison to the preceding topic/subject in discourse. The principle of coding the participants is as follows: If the topic/subject of the utterance is the same as in the preceding utterance, such a topic is not overtly marked. If there were two participants in the preceding utterance(s), the change of topic to a participant other than the one that was the topic of the previous utterance is marked through deployment of a pronoun. The pattern of coding reference of participants in propositions is as follows.

Step 1: Introduction of a new participant (participant A) through the overt mention of a noun.

Step 2: If the same participant is the only participant in the next clause, that participant is not overtly mentioned.

Step 3. If a new participant (participant B) is added, that participant is overtly coded through a noun.

Step 4. If in the next clause a reference is to be made to participant B, that reference is made through the use of a pronoun.

Here is an illustration of the steps involved. In the following fragment from Slava's narrative, in the first utterance (56) the speaker is introducing a new participant, *'par'en' tʒɛ na v'ir(ə)s' i'b'edə* “a fellow also on a bike”:

- (56) *t'ib'erə sti'f'ael ɪpar'en' tʒɛ na*
 now meet:3SG:PRES fellow:NOM also PREP:on
v'ir(ə)s'ib'edə//
 bicycle:LOC
 “Now he is meeting a fellow also on a bicycle.”

In the next utterance (57), the same topic, i.e., the fellow on the bicycle, is unmarked:

Slava

- (57) *ras// zap'ral/ u n'ivɔ/ ʃl'abu/*
 PUNCT take away:3SG:PST PREP:at 3SG:GEN hat:ACC
 “Suddenly he₁ took his₂ hat.”

In the first clause of the next utterance, the topic is marked by the pronoun *ən* “3SG.M,” which refers to the second participant of the event referred to in the preceding utterance, i.e., the fellow whose hat has been snatched. In the second clause of this utterance the topic is again unmarked, which indicates that the topic is the same as in the preceding clause, i.e., the fellow whose hat was snatched:

Slava

- (58) *i ən// ras natʒad ʔgɔlu ɪsmɔtr'i/*
 CONJ:i 3SG PUNCT back head:ACC look

Deixis

The Russian independent demonstrative *'eta* “this” has been recorded as the only deictic marker for entities (as opposed to locations) in Sino-Russian idiolects. Unlike in Russian, this marker is used to point at entity or entities regardless of the gender of the entity, the number of entities, and, most important, regardless of the distance of the entity in relationship to the speaker, to the listener, or both:

Lida

In the following example the vendor points to an article for sale:

- (59) *ʃtɔ// ɪɛdal ɪtv'esid'i p'd'is'al//*
 what DEM two.hundred fifty
 “What? This [costs] two hundred and fifty [rubles].”

Slava

Pointing at the pears in the Pear story video:

- (60) *ɪeta/ ɪknə// ɪkruʃa//*
 DEM 3PL:POSS pear:NOM
 “These are their pears.”

Egor

Referring to an event shown in the Pear story video:

- (61) *stɔ 'eta//*
 what DEM
 “What's this?”

In the recorded texts there are no instantiations of the deictic marker determining a noun, i.e., corresponding to English “this X” or “that X.”

Coding an Unknown Member of a Set

In a few idiolects there has emerged the coding of a membership in a set. This function is coded by forms derived from the Russian numeral *adin* “one” preceding the noun. The evidence that the function of the numeral is to code an unknown member in a set, rather than a single participant, is provided by the fact that the numeral *at'iin* “one” is used when the number of participants is not in question. In the following utterance relating an event in a Pear story video, the speaker uses the numeral “one” before the noun *krist'an'e* “peasant,” even though the issue of number is not in question in the utterance:

Konstantin

- (62) *ə ja 'tsz'es' ɪvit'i.ɪə/ at'iin/ krist'an'e//*
 eh 1SG here see:IPFV:PST one peasant
 “I saw a peasant here.” (“here” refers to the Pear story video).

The presence of this function may be an original creation by the speakers or may well be a copy of the function that is also encoded by the equivalent of numeral “one” in both Mandarin Chinese and in Russian, the two languages in contact for the Sino-Russian speakers. Given that this function has been observed in only a

few idiolects and in only a few utterances, this function does not interact with other functions encoded in the reference system.

Conclusions About Sino-Russian Idiolects

Many Sino-Russian idiolects code four fundamental distinctions within the reference systems: new participant, coded by the use of a lexical item; the same participant in the same role, coded by the absence of the lexical item or pronoun; switch reference, coded by use of pronouns; and deixis to entities, coded by the demonstrative *‘eta* “this, that.”

COMPARING THE COMPLEXITIES

Computing Complexity

Comparing the complexities, even within systems that have the same communicative function across languages, is a difficult proposition given the fact that even though the systems have the same communicative area within their respective languages, the functions within each system are quite different. Within the theoretical approach assumed in the present study, this is actually what is expected: There is no *a priori* reason why functions encoded in the grammatical systems across languages should be similar [see also discussions in Sampson et al. (2009), which, however, are not couched in the terms of the present approach].

One can, however, conduct the comparison of complexities in the sense of the organization of the internal system and in the number of functions a speaker of a given language has to attend to while encoding a reference in a proposition. Moreover, recall that such computation must not include functions that affect the choice of forms for the system of reference, such as the type of predication, interaction with the grammatical and semantic roles of noun phrases in the proposition, the role of the speaker, and other functions. Admittedly, this rough calculation is not very informative, as it does not take into consideration the fact that the functions through which the identity of the referent is computed in each language are different.

The following are the results of the very rough calculations of the functions that the speaker must take into consideration in the coding of reference in an utterance involving the few languages discussed in this study (each Sino-Russian idiolect constitutes an independent system). The number after the language name indicates the number of functions within the system of reference.

English has two different functions for subject as opposed to object, one for object as opposed to subject, and six different functions for identification of the noun phrase.

Mina has seven different functions for the identification of the participant in the proposition.

Polish has four different functions to identify the participants in the proposition, and five functions to identify the head of the noun phrase.

Mandarin Chinese has six functions through which the listener can identify the participants in a proposition.

Most Sino-Russian idiolects distinguish between four functions.

The results of this short study are surprising in that for the four languages that are inherited from generation to generation, namely English, Mina, Polish, and Mandarin Chinese, the number of functional distinctions within the system of reference to entities ranges between six and nine functions. One would expect these numbers would vary more because there is no theoretical limit for the number of functions to be coded within one system. For young languages, i.e., languages now being formed by adult speakers, the number of distinctions is significantly smaller.

The results of this short sample may appear to confirm what has been assumed by other scholars looking at issues of complexity, namely that the richer the morphological coding in the language, the greater the complexity. Here it is necessary to exercise caution with respect to attributing a cause-effect relationship between the function and the form. There is also evidence that the existence of coding means may be a result of the need to code a function. Thus, the elaborate logophoric system in Mupun motivates the existence of three sets of logophoric pronouns, one for the category subject, another for the category object, and a third one for other grammatical relations (Frajzyngier, 1993). Each set in Mupun codes a distinction between masculine singular, feminine singular, and plural pronouns. The presence of the rich set of pronouns is driven by the functions coded in the grammatical system. The relationship between form and function, the basis of any complexity in the grammatical system, is therefore a bidirectional relationship in which either the form or the meaning could be either the cause or the effect.

One of the questions with which this study started is what the notion of complexity in the grammatical system is good for. The study asserts that the whole-language complexity has no heuristic value. Even if somebody proposes a metric for the whole-language complexity it is not clear what such a metric can be used for. On the other hand, a metric of complexity within a given functional domain has several theoretical and practical applications.

Practical applications are those that have always faced the practical applications of linguistics. First-language acquisition studies in the domain of phonology have demonstrated long ago that the acquisition of a complex phonological system, i.e., a system with a larger number of underlying segments and a large number of rules of their realization, takes longer than the acquisition of the phonological system with a smaller number of segments and smaller number of rules of realization. We do not have comparable studies of the acquisition of the totality of semantic structure encoded in a language, because no such goal has been set up by researchers.

Second-language acquisition demonstrates that acquiring a functional domain in L2 which is more complex than a similar domain in L1 is more difficult than acquiring a simpler system, i.e., a system with fewer semantic distinctions. Thus, acquiring a gender system in L2 when L1 has no gender system often results in a haphazard assignment of gender by L1 speakers speaking L2.

Complexity also plays a role in language loss for multi-lingual speakers when they shift to another language and for mono-lingual speakers under language impairment. The common thread appears to be the reduction of complexity in some functional domains. There are more questions here than answers. For example, which meanings are lost first, and which meanings are lost later? In order to answer this and other questions one needs to have an explicit description of the complexity of the given domain. The complexity of any functional domain changes over time, thus supporting Sampson (2009) and other studies in Sampson (2009).

The explicit understanding of complexity within a given functional domain is a crucial prerequisite for the analysis of the functions in a language and for linguistic typology. The cross-linguistic studies centered on some “prototypical” or “canonical” definitions of functions, e.g., “indefinite,” “definite,” “perfective,” or “future,” or “singular,” are bound to be of limited value or even misleading, if they do not consider the complexity of the functional domain to which the given function belongs. If one ignores the complexity of the domain, one in fact does not compare the meanings/functions of the forms under study but rather what motivated a given linguist to assign one label, rather than another, to a given form. This would be similar to comparing the sign “3” on a clock that has 24-h division with number “3” on a clock that has 12-h division. In order to understand any function/meaning encoded in the grammatical system, one needs to know what other functions are encoded in the given domain. Complexity of a functional domain is a necessary factor to be taken into consideration in the discovery and the description of the individual functions.

REFERENCES

- Carlson, G. N. (1980). *Reference to Kinds in English*. New York, NY: Garland.
- Corrie, B. (1998). Reference tracking: description and explanation. *Sprachtypol. Universalienforsch.* 51, 335–346.
- Corbett, G. G. (2006). *Agreement*. Cambridge: Cambridge University Press.
- Dahl, Ö. (2004). *The Growth and Maintenance of Linguistic Complexity*. Amsterdam, Philadelphia, PA: John Benjamins.
- de Swart, H., and Zwarts, J. (2009). Less form – more meaning: why bare singular nouns are special. *Lingua* 119, 280–295. doi: 10.1016/j.lingua.2007.10.015
- Delfitto, D. (2006). “Bare plurals,” in *The Blackwell Companion to Syntax*, eds M. Everaert and H. van Riemsdijk (Malden, MA: Blackwell Publishing). doi: 10.1002/9780470996591.ch8
- Dixon, R.M.W. (2010). *Basic Linguistic Theory. Vol. 1: Methodology*. Oxford: Oxford University Press.
- Dixon, R. M. W. (2016). *Are Some Languages Better Than Others?* Oxford: Oxford University Press.
- Frajzyngier, Z. (1985). Logophoric systems in Chadic. *J. Afr. Lang. Linguist.* 7, 23–37.
- Frajzyngier, Z. (1991). “The de dicto domain in language,” in *Approaches to Grammaticalization, Vol. 1*, eds E. C. Traugott and B. Heine (Amsterdam, Philadelphia, PA: John Benjamins), 219–251.
- Frajzyngier, Z. (1993). *A Grammar of Mupun*. Berlin: Reimer.
- Frajzyngier, Z. (1997). “Pronouns and agreement: systems interaction in the coding of reference,” in *Atomism and Binding*, eds H. Benis, P. Pica, and J. Rooryck (Dordrecht: Foris), 115–140.
- Frajzyngier, Z. (2001). *A Grammar of Lele*. Stanford, CA: CSLI.
- Frajzyngier, Z. (2019). An integrated approach to lexicon, syntax, and functions. *J. Linguist. Soc. N. Zeal.* 62, 1–23.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

Work on selected topics in this paper was supported by a University of Colorado Small Grant program from the Center for Humanities & the Arts.

ACKNOWLEDGMENTS

I am grateful to my wife, Anna, who proposed that the Polish suffix *-no* actually represents the event itself, rather than who did what. I am grateful to three anonymous readers of an earlier version of this study. Marcin Kilarski went through the text with a sharp eye for all aspects of this study. His suggestions improved the papers in several ways. I am also grateful to Erin Shay for her careful reading, and for substantial and editorial comments on this paper. The readers’ comments saved me from one or two embarrassing mistakes and prompted me to consider issues not tackled in the previous versions. All errors of fact or interpretation are my sole responsibility.

- Frajzyngier, Z., and Butters, M. (2020). *The Emergence of Grammatical Functions*. Oxford: Oxford University Press.
- Frajzyngier, Z., Gurian, N., and Karpenko, S. (2021). *Grammar Formation by Adults: The Case of Sino-Russian Idiolects*. Leiden, Boston, MA: Brill.
- Frajzyngier, Z., and Johnston, E., with Edwards, A. (2005). *A Grammar of Mina*. Berlin, New York, NY: Mouton de Gruyter.
- Frajzyngier, Z., Liu, M., and Ye, Y. (2020). The reference system of Modern Mandarin. *Aust. J. Linguist.* 40. doi: 10.1080/07268602.2019.1698512
- Frajzyngier, Z., and Shay, E. (2003). *Explaining Language Structure Through Systems Interaction*. Amsterdam, Philadelphia, PA: John Benjamins.
- Frajzyngier, Z., with Shay, E. (2016). *The Role of Functions in Syntax: A Unified Approach to Language Theory, Description, and Typology*. Amsterdam, Philadelphia, PA: John Benjamins.
- Gragg, G., and Hoberman, R. (2012). “Semitic,” in *The Afroasiatic Languages*, eds Z. Frajzyngier and E. Shay (Cambridge: Cambridge University Press), 145–235.
- Gundel, J., and Abbott, B. (2019). (eds.). *The Oxford Handbook of Reference*. Oxford: Oxford University Press.
- Laskowski, R. (1984). “Zaimek (Pronoun),” in *Gramatyka Współczesnego Języka Polskiego. Morfologia. (Grammar of contemporary Polish. Morphology.)*, eds R. Grzegorzczkowska, R. Laskowski, and H. Wróbel (Warsaw: Państwowe Wydawnictwo Naukowe), 275–282.
- Lazard, G. (2004). On the status of linguistics with particular regard to typology. *Linguist. Rev.* 21, 389–411. doi: 10.1515/tlr.2004.21.3-4.389
- Le Bruyn, B., de Swart, H., and Zwarts, J. (2017). *Bare nominals. Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. doi: 10.1093/acrefore/9780199384655.013.399
- Longobardi, G. (2001). How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Nat. Lang. Seman.* 9, 335–369. doi: 10.1023/A:1014861111123

- Matthews, P. (1997). *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- McWhorter, J. (2001a). The world's simplest grammars are creole grammars. *Linguist. Typol.* 5, 125–166. doi: 10.1515/lity.2001.001
- McWhorter, J. (2001b). What people ask David Gil and why: rejoinder to the replies. *Linguist. Typol.* 5, 388–412. doi: 10.1515/lity.2001.003
- McWhorter, J. (2009). “Oh nó!: a bewilderingly multifunctional Saramaccan word teaches us how a creole language develops complexity,” in *Sampson, Gil, and Trudgill*, 141–163.
- Newmeyer, F., and Joseph, J. (2012). “All languages are equally complex”: the rise and fall of a consensus. *Historiogr. Linguist.* 39, 341–368. doi: 10.1075/hl.39.2-3.08jos
- Nitsch, K. (1960). *Wybór Polskich Tekstów Gwarowych. (A Selection of Polish Dialect Texts)*. Warsaw: Państwowe Wydawnictwo Naukowe.
- NKJP. Narodowy Korpus Języka Polskiego. *Polish National Corpus*. Available online at: <http://nkjp.pl/poliqarp/nkjp300/query/> (Consulted at various times)
- Payne, J., and Huddleston, R. (2002). “Nouns and noun phrases,” in *The Cambridge Grammar of the English Language*, eds R. Huddleston and G. K. Pullum (Cambridge: Cambridge University Press), 323–524.
- Quirk, R., and Greenbaum, S. (1973). *A Concise Grammar of Contemporary English*. New York, NY: Harcourt Brace Jovanovich.
- Roberts, I. (2019). *Parameter Hierarchies and Universal Grammar*. Oxford: Oxford University Press.
- Sampson, G. (2009). “A linguistic axiom challenged,” in *Sampson, Gil, and Trudgill*, 2–18.
- Sampson, G., Gil, D., and Trudgill, P. (2009). (eds.). *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.
- Stvan, L. S. (2007). “The functional range of bare singular count nouns in English,” in *Nominal Determination. Typology, Context Constraints, and Historical Emergence*, eds E. Stark, E. Leiss, and W. Abraham (Amsterdam, Philadelphia, PA: John Benjamins), 171–187.
- Topolińska, Z. (1984). “Składnia grupy imiennej. (Syntax of the noun phrase),” in *Gramatyka Współczesnego Języka Polskiego. Morfologia*, eds R. Grzegorzczkowska, R. Laskowski, and H. Wróbel (Warsaw: Państwowe Wydawnictwo Naukowe), 301–386.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Frajzyngier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.