Maria Irena Szawerna

Lab 1: Chapter 2 report

<u>NOUNS</u>

In Polish, nouns are inflected for case and number, and have the inherent feature of gender (also known as class sometimes). There exist both simple and compound nouns, and many distinctions (both grammatical and semantic) can be made within this category, however, only some of them are relevant for this assignment.

Overall, there are two numbers, with leftovers of a third one. Most nouns have a singular and plural form, and some have a dual one as well. There are mass nouns and abstract nouns that rarely have a plural form. There are also nouns that only occur in plural.

Cases in Polish influence nouns, adjectives pronouns, numerals, and adjectival participles. Polish has 7 different cases: nominative, accusative, genitive, dative, locative, instrumental, vocative. The nominative is the form used for subjects of a sentence and is the basic form of a noun. Accusative is the case for the direct object. Dative is the case for indirect objects and logical subjects of impersonal verbs. Locative is the case for pointing to a location (figurative or literal). Instrumental is the case which signifies that the noun is being used as a tool or means of some action. Vocative is the case used when addressing someone. Many of the cases can be triggered by specific verbs of prepositions that they are required by (rection).

In Polish we can identify three genders: masculine, feminine, and neuter. Masculine is further divided into animate and inanimate. There is some regularity to how this is assigned, namely on the basis of the biological gender or gender identity of the referent or the nominative case ending of the word. In the plural, the divisions are slightly different. There is the plural masculine human, plural masculine non-human, plural feminine and plural neuter; the non-human plurals are often grouped together, e.g. in the case of inflecting verbs by gender.

Declension patterns do not necessarily align with the gender of the noun, e.g. here are singular masculine nouns that can follow the feminine declension pattern, but the adjectives, verbs, etc., will have masculine endings. This is because in general the declension depends on the structure of the noun (i.e. what sounds it ends in), and there are masculine nouns that look like feminine ones (they end with -a). Those nouns thus take the feminine declension pattern but in terms of agreement they require adjectives, verbs, etc. to take masculine forms, so the gender of the noun is masculine.

There is some regional, personal, or social variation when it comes to some inflectional endings. The inflectional paradigms are decided based on the gender of the noun and the

structure of its stem. Overall, for nouns of all numbers and genders the following declension paradigms can be identified:

1. Stems ending in a velar consonant (k, g, ch, h)
2. Stems ending in a "historically soft" consonant (c, dz, sz, dz, ż, cz, dż)
3. Stems ending in a soft consonant (ś, ć, ź, dź, ń, l, soft b, soft m, soft w, soft p, j)
4. Stems ending in all other consonants – hard stems

Nevertheless, there is variation within those, and there are special cases and/or exceptions from these rules.

There exist pejorative, augmentative, and diminutive suffixes for nouns, which, together with abbreviations, have their rules for assigning gender and declension. There are nouns where the gender can differ depending on "personal preference"; there are also homographs (and homophones) that have different meaning depending on the gender.

Nouns must agree with the following: adjectives (number, gender, case), demonstrative pronouns (same), personal pronouns (same), 1st and 2nd person singular and plural possessive pronouns (same), singular past tense verb forms (number and gender), future progressive tense (same), conditional verb forms (same), adjectival participles (number, gender, case) (Sadowska, 2012, pp. 40-217).

ADJECTIVES

As for adjectives, they inflect for number, case, and gender of the noun they describe. They also have the base, comparative, and superlative forms. The cases, numbers, and genders are the same as described for nouns in the previous section. The differences in declension result from the same stem differences as in nouns. However, there is a smaller variation of endings, and many case forms take the same endings. Singular masculine adjectives end in -i (k, g, soft stem) or -y. Feminine ones always end in -a, and neuter ones in -e or -ie (k, g, soft stem). In plural male human adjectives end in -y or -i, and the rest in -e. Stem alterations can happen. Adjectives can occur in pre- or post-nominal position; the first option describes some quality of the noun, while the second one is used to classify it (e.g. "brunatny niedźwiedź", literally "brown bear", means a bear that is brown, while "niedźwiedź brunatny", literally "bear brown", means a bear belonging to the species called the brown bear). Some adjectives have short variants, but this is slightly archaic and not productive.

Some adjectives can impose case on the following noun. This happens in cases where in English we would need some preposition, e.g. "unworthy **of** trust" or "close **to** heart". There are also certain feminine nouns ending in -owa are declined like adjectives. Negated adjectives are written as one word, but adverbs are written mostly separately. Some intensifiers are written

together, and some separately. There are also suffixes that signify emotive versions of adjectives.

Comparative forms are created by adding a suffix -szy/-ejszy in masculine (and the last vowel changes depending on the gender of the noun described). Superlative forms are created by adding a prefix naj- and the same suffix as in comparative, while lower and lowest degree are created using adverbs "mniej" (less) and "najmniej" (the least). Some adjectives have completely irregular comparative and superlative forms, and other adjectives yet do not have comparative and superlative forms and instead use adverbs such as "bardziej" (more) and "najbardziej" (the most).

There exist demonstrative adjectives, which are roughly the same as English determiners (this, that, etc.). They have their own declensional paradigm. It is also the case for distributive adjectives (each, all), as well as the adjectives meaning "some" and "other". "Który", "jaki", "czyje" (which, what (kind of), whose) are also considered adjectives, so-called interrogative adjectives and they need to agree with the noun they modify. There also exist relative adjectives which are declined like regular adjectives. It is peculiar since these would perform the roles of what is described as determiners and pronouns in English, but it just points to how unusual these words are (Sadowska, 2012, pp. 218-264).

<u>VERBS</u>

Polish verbs conjugate to reflect the person, number, tense, mood, voice, aspect, and sometimes gender of the subject. Personal pronouns are often omitted since all the necessary information is conveyed by the verb. As for the number, singular and plural are distinguished. As for genders, masculine, feminine, neuter singular and human male and no human male plurals are relevant. Not all of the genders are marked in every tense: they are relevant for every number and person only in constructions using past tense verb forms.

The two aspects in Polish are the perfective and imperfective one. Verbs appear in pairs that represent the perfective and imperfective aspects. The pairs are created by adding a prefix to the imperfective infinitive, shortening or modifying the stem of the imperfective infinitive, using two completely different verbs. One imperfective form can have many perfective counterparts that may vary in semantics (e.g., to have been eating vs. to have eaten vs. to have eaten up vs. to have eaten a bit, vs. to have eaten enough not to be hungry); usually the meaning can be partly guessed from the prefix used. Depending on the aspect, the verbs can only take some tenses: perfective aspect does not occur in the present tense. Some verbs can serve both as the perfective and imperfective form, and some appear only in perfective or only

3

in imperfective. Some verbs have two versions of the imperfective aspect, one of which being the regular one, and the other the habitual or frequentative or iterative verbs.

There are four conjugation patterns for Polish verbs: I. verbs ending in -ać, II. verbs ending in -eć, III. diverse verbs, including pseudo-ać, IV. verbs ending in -yć/-ić/-eć.

The following moods can be distinguished: indicative, imperative, conditional, subjunctive. Not all moods can occur in all tenses or for all persons. The following tenses can be distinguished: (infinitive), present, past, (archaic) past perfect (formed by "to be" + past form), future simple (perfective – using the perfective form of the verb in the present tense form), future compound (imperfective – using the future form of "to be" with the imperfective infinitive). A special set of endings (also inflected by number, person, gender) are used to create conditional forms. They can be used as suffixes, but also detached and used as particles separately. Imperative forms are based off of the infinitive. The ending thereof is dropped, and sometimes stem sound changes occur. For the 2nd person singular that is the form, while for second person plural and first person plural suffixes are added (-cie, -my).

There are three adjectival participles (present active, present passive, past passive) and two adverbial participles (present and past). All of those can be formed both from perfective and imperfective forms. The adjectival ones are declined using adjectival paradigms. The adverbial ones are not declined. There exist also impersonal forms of personal verbs (somewhat similar to passive ones).

There are two voices: active and passive. The passive voice is formed by using auxiliary verbs "być/to be" (for imperfective) or "zostać/to become" (for perfective) and the past participle. The participle agrees with the person/thing acted upon (so the subject of the newly formed sentence). To mark the performer of the action the preposition "przez" and the accusative case or just the accusative case can be used.

Reflexive verbs take the suffix "się" (spelled separately) in all the forms.

The verb "być" (to be) is mostly used as a linking/copula verb, but is also sometimes an auxiliary (compound future tense, passive voice, modal verbs, impersonal expressions). Sometimes it can be used independently (or with an implied dummy subject). When negated, "być" only inflects by tense. Modal verbs are also inflected. They are followed by infinitives. There are also modals that take no personal endings and can only be used with unspecified subjects.

There exist impersonal verbs that do not have personal forms and are used as the infinitive or third person singular (neuter wherever applicable). These most often refer to weather conditions, but not only (Sadowska, 2012, pp. 306-462).

To sum up, in Polish the following holds true:

- Nouns have the following features:
  - Inherent: gender
  - Inflectional: case, number
- Adjectives have the following features:
  - Inherent: length
  - Inflectional: case, number, gender, degree
- Verbs have the following features:
  - Inherent: aspect, reflexivity
  - Inflectional: person, number, tense, mood, voice, gender

These features can take the following values:

- Gender: masculine, feminine, neuter in singular, masculine human, masculine non-human, feminine, neuter in plural (the last three are often grouped together as the distinction often does not exist).
- Case: nominative, accusative, genitive, dative, locative, instrumental, vocative.
- Number: singular, plural, archaic dual.
- Length: short, long.
- Degree: positive, comparative, superlative.
- Aspect: perfective, imperfective.
- Reflexivity: reflexive, not reflexive.
- Person: 1st, 2nd, 3rd , impersonal.
- Tense: present, past, archaic past perfect, future simple, future compound.
- Mood: indicative, imperative, conditional, subjunctive.
- Voice: active, passive.

In addition, conjugations and declensions are conditioned by the sounds found in the stem of the word.

<center>COMPARISON</center>

In order to compare the findings from the previous section with the PUD treebank annotation for Polish, I extended the functionality of the [Python script](#) that I wrote for the Chapter 1 lab. I could now use it to retrieve the features tags sorted by popularity, both when kept in their original combinations (e.g. ('Case=Loc', 'Degree=Sup', 'Gender=Neut', 'Number=Sing')) or one by one, using the --feats and --indi_feats flags. I also now did not have to put in a number for the top X results, which let me easily print all of them without guessing how many unique ones there are.

In the Polish PUD treebank there are 668 different combinations of features present. These are composed of a total of 49 individual feature tags. Looking at the individual tags I found the following that corresponded to my description of the grammar: Number=Sing, Gender=Masc, Gender=Fem, Animacy=Inan, Case=Gen, Number=Plur, Case=Nom, Degree=Pos, Voice=Act, Gender=Neut, Mood=Ind, Case=Acc, Aspect=Imp, Case=Loc, Variant=Short, Aspect=Perf, Tense=Past, Person=3, Tense=Pres, Case=Ins, Reflex=Yes, Voice=Pass, Variant=Long, Case=Dat, Person=1, Degree=Cmp, Tense=Fut, Degree=Sup, Person=0, Person=2, Mood=Imp, Case=Voc. As well as the following that I seem to have missed or described differently: VerbForm=Fin, Animacy=Hum, Polarity=Pos, VerbForm=Part, VerbForm=Inf, VerbForm=Vnoun, Polarity=Neg, Abbr=Yes, Animacy=Nhum, Poss=Yes, VerbForm=Conv, VerbType=Quasi, VerbType=Mod.

From the ones that I have missed one of the more important categories is the verb form, which describes whether the verb is finite, in the infinitive, a particle, or a noun derived from a verb. Then animacy was also split off from gender, and is described separately as human, non-human, inanimate. Verb type indicates if it is a modal verb or some other special kind of verb. Finally polarity is a feature given to participles to denote whether they are negated or not (negation would be in the form of a prefix).

In comparison with my list, the dual number is missing, but this may be due to there being very few nouns that take that number, and those are not extremely likely to appear in this text (they refer to eyes, arms, and ears). As mentioned before, gender was approached slightly differently. Impersonal verbs were given a Person=0 tag. Tags for compound tenses were not specified (compound future, past perfect) as these consist of more than one word and each of those words can be tagged differently. Conditional and subjunctive mood did not appear in the tags, perhaps due to lack of them in the source text. Everything else aligned with my findings. What is perhaps worth noting is that within these features here there is nothing to suggest which

declension to use. Perhaps it is because in many cases it is possible to guess that from the ending of the stem.

Hopefully this comparison is sufficient and the description of features of nouns, verbs, and adjectives is not too voluminous; sadly, it was difficult to shorten it any more, as Polish is a highly inflected language and there are many features that are conveyed in that fashion.

References:

Sadowska, I. (2012). *Polish: A comprehensive grammar*. Routledge.