Morphology IV: Capturing Syncretism

Underspecification, Decomposition & The Subset Principle

Philipp Weisser

Universität Leipzig
philipp.weisser@uni-leipzig.de

14. Dezember 2018

- In this set of slides, we will get to know some of the basic concepts shared by many (though not all) formal morphological theories.
  - Underspecification
  - The Subset Principle
  - Feature Decomposition
  - Feature Geometries
  - Featural Markedness

- Even though these concepts are often used together, it must be noted that they are formally independent.
  - ↪ Some theories assume Underspecification but do not assume the Subset Principle
  - ↪ Some theories reject the notion of decomposition and/or feature geometries.
  - ↪ The concept of markedness is often controversial even within a given framework

- As with all building blocks of morphological theories, the main motivation for these concepts comes from mismatches between syntax and morphology.

- As with all building blocks of morphological theories, the main motivation for these concepts comes from mismatches between syntax and morphology.
- In particular syncretism, a one-to-many mapping between syntactic context and morphological exponent will be the driving force behind many of the assumptions in this set of slides.

- As with all building blocks of morphological theories, the main motivation for these concepts comes from mismatches between syntax and morphology.
- In particular syncretism, a one-to-many mapping between syntactic context and morphological exponent will be the driving force behind many of the assumptions in this set of slides.
- However, the concepts of underspecification, decomposition, markedness, etc. are often also useful for describing other mismatches (allomorphy, multiple exponence, etc.)

Underspecification

- Languages differ as to whether they have a designated morphological form for specific syntactic context.

- Languages differ as to whether they have a designated morphological form for specific syntactic context.
- In some (often agglutinative) languages, all (or at least most) syntactic contexts are faithfully mapped onto morphological forms.
  - ↪ The syntactic context unambiguously predicts the morphological form
  - ↪ And the morphological form unambiguously encodes the syntactic context.

- Consider the following paradigm of Finnish present tense verbal markers:

(1) Finnish Present Tense

1sg	puhu-t
2sg	puhu-n
3sg	puhu-∅
1pl	puhu-mme
2pl	puhu-tte
3pl	puhu-vat

- Each syntactic context (1sg, 3pl, etc.) corresponds to exactly one morphological marker.

- More often than not, languages do not employ as many morphological markers as there are syntactic contexts.

- More often than not, languages do not employ as many morphological markers as there are syntactic contexts.
- Often, a number of syntactic contexts are expressed with the same morphological form/exponent. This mismatch is called *syncretism*.

- More often than not, languages do not employ as many morphological markers as there are syntactic contexts.
- Often, a number of syntactic contexts are expressed with the same morphological form/exponent. This mismatch is called *syncretism*.
- But typically, languages do not seem to collapse random sets of syntactic contexts to yield one and the same form. Some syncretism patterns are more common than others.

- In some cases, certain morphological categorizes receive the same form.

(2) Palatinate German Present tense of *lavre* (speak):

1sg	lav-ɐ
2sg	lav-ɐʃ
3sg	lav-ɐt
1pl	lav-ɐn
2pl	lav-ɐn
3pl	lav-ɐn

- In the case of this paradigm, the plural receives the same marker for all three persons.
- We say that the feature person was *neutralized* in the context of the feature plural.

- As alluded to in the last set of slides, patterns of this sort indicate that morphological forms are in fact *underspecified* for their syntactic context.
  - ↪ They are compatible with more than one combination of morphosyntactic features.

- As alluded to in the last set of slides, patterns of this sort indicate that morphological forms are in fact *underspecified* for their syntactic context.
  - ↪ They are compatible with more than one combination of morphosyntactic features.

Underspecification

A given morphological marker/exponent can realize/contribute only a subset of the morphosyntactic features present in the syntax.

- In a realizational model, the syntax is fully specified for morphosyntactic features but the morphology only realizes a subset of these.

- In a realizational model, the syntax is fully specified for morphosyntactic features but the morphology only realizes a subset of these.
- In an incremental model, the affixes combined in word formation contribute only a subset of features that are present in the syntax. Since only the features of individual words must be checked (and not the features on functional heads), the derivation still converges.

- In a realizational model, the syntax is fully specified for morphosyntactic features but the morphology only realizes a subset of these.
- In an incremental model, the affixes combined in word formation contribute only a subset of features that are present in the syntax. Since only the features of individual words must be checked (and not the features on functional heads), the derivation still converges.
  - ↪ Recall: This distinction was the basis of our argument for incremental models last week.

- Different frameworks have different ways to notate underspecified items but the intuition is the same. Not all affixes are specified for all types of features.

- Different frameworks have different ways to notate underspecified items but the intuition is the same. Not all affixes are specified for all types of features.
- In the case of the Palatinate German verbal inflection syncretism, all affixes are specified for person, number and tense except for the one in (3-d). It does not have a person specification.

(3)
a. [1.SG.Pres] ↔ /ɐ/
b. [2.SG.Pres] ↔ /ɐʃ/
c. [3.SG.Pres] ↔ /ɐt/
d. [PL.Pres] ↔ /ɐn/

- Different frameworks have different ways to notate underspecified items but the intuition is the same. Not all affixes are specified for all types of features.
- In the case of the Palatinate German verbal inflection syncretism, all affixes are specified for person, number and tense except for the one in (3-d). It does not have a person specification.

(3)
a. [1.SG.Pres] ↔ /ɐ/
b. [2.SG.Pres] ↔ /ɐʃ/
c. [3.SG.Pres] ↔ /ɐt/
d. [PL.Pres] ↔ /ɐn/

- (3-d) is *underspecified* for person.

- Unless underspecification is supplemented with feature geometries or a theory of markedness (see slides below), a marker could also be specified for person and underspecified for number:

(4) Kuman (Trans-New Guinea) Pronouns:

	SG	PL
1	na	no
2	ene	ene
3	ye	ye

↪ Here, second and third person are underspecified for number.

- Abstractly speaking, underspecification is a means to model syncretism of adjacent cells in a paradigm forming a *natural class*:
  - ↪ Natural Classes are sets of morphosyntactic contexts defined by the common presence of a certain feature:

(5)

	A	B
$-\alpha$	♣	♠
$+\alpha$	♦	♦

The Subset Principle

- Underspecification often is not enough to model syncretism patterns in the world's languages.

- Underspecification often is not enough to model syncretism patterns in the world's languages.
- Often, it fails to derive *default effects*.

(6) Englisches Present tense paradigm:

1sg	walk-∅
2sg	walk-∅
3sg	walk-s
1pl	walk-∅
2pl	walk-∅
3pl	walk-∅

- Underspecification often is not enough to model syncretism patterns in the world's languages.
- Often, it fails to derive *default effects*.

(6) Englisches Present tense paradigm:

1sg	walk-∅
2sg	walk-∅
3sg	walk-s
1pl	walk-∅
2pl	walk-∅
3pl	walk-∅

- Here, it is not simply a whole category that is neutralized. Rather, it seems that there is one fully specified marker and then a second *default* marker that is used when the other one is not applicable.

- The assumption of underspecification can have the effect that several markers are potentially applicable to the same morphosyntactic context.

- The assumption of underspecification can have the effect that several markers are potentially applicable to the same morphosyntactic context.
  - ↪ This competition is resolved by the insertion contexts. More specified markers take precedence over less specified markers.

- The assumption of underspecification can have the effect that several markers are potentially applicable to the same morphosyntactic context.
  - ↪ This competition is resolved by the insertion contexts. More specified markers take precedence over less specified markers.
- This models the intuition that a default marker is always applicable unless it is blocked by a more specific one.

- The assumption of underspecification can have the effect that several markers are potentially applicable to the same morphosyntactic context.
  - ↪ This competition is resolved by the insertion contexts. More specified markers take precedence over less specified markers.
- This models the intuition that a default marker is always applicable unless it is blocked by a more specific one.
- Formally, this is achieved by the Subset Principle and the definition of Specificity.

Subset Principle

A Vocabulary Item V is inserted into a syntactic head X iff (i) and (ii) hold:

(i) The morphosyntactic features of V are a subset of the morphosyntactic features of X.

(ii) V is the most specific Vocabulary Item fulfilling (i)

Subset Principle

A Vocabulary Item V is inserted into a syntactic head X iff (i) and (ii) hold:

(i) The morphosyntactic features of V are a subset of the morphosyntactic features of X.

(ii) V is the most specific Vocabulary Item fulfilling (i)

Specificity

A Vocabulary Item V_i is more specific than a Vocabulary Item V_j iff it has more features.

- This implementation allows us to derive a great amount of syncretism
  - ↪ Following our intuition earlier, the English verbal inflection paradigm can simply be modelled using two Vocabulary Items:

(7) a. $[_T$,3.SG] ↔ /-s/
b. $[_T\ ]$ ↔ /-∅/

- This implementation allows us to derive a great amount of syncretism
  - ↪ Following our intuition earlier, the English verbal inflection paradigm can simply be modelled using two Vocabulary Items:

(7) a. $[_T$,3.SG] ↔ /-s/
b. $[_T\ ]$ ↔ /-∅/

- (7-a) is fully specified for person and number (ignoring gender for now). (7-b) is radically underspecified. It realizes no features at all.

- This implementation allows us to derive a great amount of syncretism
  - ↪ Following our intuition earlier, the English verbal inflection paradigm can simply be modelled using two Vocabulary Items:

(7) a. $[_T$,3.SG] ↔ /-s/
b. $[_T\ ]$ ↔ /-∅/

- (7-a) is fully specified for person and number (ignoring gender for now). (7-b) is radically underspecified. It realizes no features at all.
- Given the feature specification, both VIs could be inserted in [3.SG]-contexts. However, since (7-a) is more specific, it takes precedence.

- This implementation allows us to derive a great amount of syncretism
  - ↪ Following our intuition earlier, the English verbal inflection paradigm can simply be modelled using two Vocabulary Items:

(7) a. $[_T$,3.SG] ↔ /-s/
b. $[_T\ ]$ ↔ /-∅/

- (7-a) is fully specified for person and number (ignoring gender for now). (7-b) is radically underspecified. It realizes no features at all.
- Given the feature specification, both VIs could be inserted in [3.SG]-contexts. However, since (7-a) is more specific, it takes precedence.
- In all other contexts, only (7-b) can be inserted because the features of (7-a) are no subset of the morphosyntactic context.

- Abstractly speaking, the combination of underspecification and the Subset Principle allows us to derive:
  - Syncretism patterns as the result of neutralization effects where morphosyntactic distinctions are not expressed in a given context.

- Abstractly speaking, the combination of underspecification and the Subset Principle allows us to derive:
  - Syncretism patterns as the result of neutralization effects where morphosyntactic distinctions are not expressed in a given context.
  - Default effects, where an underspecified marker (potentially even radically underspecified) occurs in a number of cells in the paradigm but is sometimes blocked by the occurence of a more specific marker:

- Abstractly speaking, the combination of underspecification and the Subset Principle allows us to derive:
  - Syncretism patterns as the result of neutralization effects where morphosyntactic distinctions are not expressed in a given context.
  - Default effects, where an underspecified marker (potentially even radically underspecified) occurs in a number of cells in the paradigm but is sometimes blocked by the occurence of a more specific marker:

(8)

	A	B	C	D
$-\alpha$	♦	♠	♦	♥
$+\alpha$	♦	♣	♦	♦

- Abstractly speaking, the combination of underspecification and the Subset Principle allows us to derive:
  - Syncretism patterns as the result of neutralization effects where morphosyntactic distinctions are not expressed in a given context.
  - Default effects, where an underspecified marker (potentially even radically underspecified) occurs in a number of cells in the paradigm but is sometimes blocked by the occurence of a more specific marker:

(8)

	A	B	C	D
$-\alpha$	♦	♠	♦	♥
$+\alpha$	♦	♣	♦	♦

(9)

a. $[-\alpha,\text{B}] \leftrightarrow$ /♠/
b. $[-\alpha,\text{D}] \leftrightarrow$ /♥/
c. $[+\alpha,\text{B}] \leftrightarrow$ /♣/
d. $[\ ] \leftrightarrow$ /♦/

- Further, we can derive what one might call relativized default effects (see Aronoff 2013).
  - ↪ Here, no marker is (or needs to be) relatively underspecified. Rather, it seems that there are several defaults for different subparadigms:

(10)

	A	B	C	D
$-\alpha$	♠	♠	♠	♥
$+\alpha$	♦	♣	♦	♦

Some notational and terminological points:

- The Subset Principle is sometimes also referred to as Pāṇini's Principle or the Elsewhere Principle or the Elsewhere Condition.

Some notational and terminological points:

- The Subset Principle is sometimes also referred to as Pāṇini's Principle or the Elsewhere Principle or the Elsewhere Condition.
- The notation throughout this illustration of Underspecification and the Subset Principle is DM-notation. But many other frameworks use the same concepts.

Some notational and terminological points:

- The Subset Principle is sometimes also referred to as Pāṇini's Principle or the Elsewhere Principle or the Elsewhere Condition.
- The notation throughout this illustration of Underspecification and the Subset Principle is DM-notation. But many other frameworks use the same concepts.
-

Decomposition

- With the tools we have developed, we can derive quite a number of abstract syncretism patterns. However, one well-attested pattern cannot yet be derived.

- With the tools we have developed, we can derive quite a number of abstract syncretism patterns. However, one well-attested pattern cannot yet be derived.
- We often find that some categories tend to cluster together more often than others.

- With the tools we have developed, we can derive quite a number of abstract syncretism patterns. However, one well-attested pattern cannot yet be derived.
- We often find that some categories tend to cluster together more often than others.
- In Germanic languages, for some reason, first and third person seem to be syncretic quite systematically:

(11) English Copula (Past)

	SG	PL
1	was	were
2	were	were
3	was	were

(12) English Copula (Past)

	SG	PL
1	was	were
2	were	were
3	was	were

- This pattern can, with current means, not be described. /were/ seems to be the default but there are two cells where it is blocked by /was/.

(12) English Copula (Past)

	SG	PL
1	was	were
2	were	were
3	was	were

- This pattern can, with current means, not be described. /were/ seems to be the default but there are two cells where it is blocked by /was/.
- The intuition that many people have is that first and third person have something in common to the exclusion of second person.

- Similar cases are often found with syncretism in case marking. Some cases seem to cluster in terms of form:

(13) Determiner inflection in German (SG):

	Mask	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- Similar cases are often found with syncretism in case marking. Some cases seem to cluster in terms of form:

(13) Determiner inflection in German (SG):

	Mask	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- Accusative and nominative often receive the same marking.
- Genitive and dative are sometimes also syncretic.

- Moreover, in this paradigm, it seems that masculine and neuter seem to have something in common (to the exclusion of feminine).

(14) Determiner inflection in German (SG):

	Mask	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- Moreover, in this paradigm, it seems that masculine and neuter seem to have something in common (to the exclusion of feminine).

(14) Determiner inflection in German (SG):

	Mask	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- How can we thus capture the intuition that some morphosyntactic categories are more closely related than others?

- The solution that has been proposed to model these kinds of patterns is called *decomposition*.
  - ↪ We assume that morphosyntactic categories (such as accusative, second person or masculine) are not primitives of the theory but can be *decomposed* into more abstract features.

- The solution that has been proposed to model these kinds of patterns is called *decomposition*.
  - ↪ We assume that morphosyntactic categories (such as accusative, second person or masculine) are not primitives of the theory but can be *decomposed* into more abstract features.
  - ↪ Commonalities across different categories can then be modelled by assuming that two categories share a more abstract feature (value).

- Decomposition simply adds one level of abstractness to what we call morphosyntactic features.

- Decomposition simply adds one level of abstractness to what we call morphosyntactic features.
- It essentially defines natural classes of morphosyntactic categories.

- Decomposition simply adds one level of abstractness to what we call morphosyntactic features.
- It essentially defines natural classes of morphosyntactic categories.
- The same is done in phonology by assuming that the classes of plosives and fricatives are members of a natural class called obstruents.

- For case marking, people have often assumed that there is an abstract feature shared by structural cases such as accusative and nominative and one which is shared by obligue cases such as dative or accusative.

- For case marking, people have often assumed that there is an abstract feature shared by structural cases such as accusative and nominative and one which is shared by obligue cases such as dative or accusative.
- And in order to do justice to the observation that dative and accusative share some kind of feature, they assumed that accusative and dative share a feature to the exclusion of nominative and genitive.

- For case marking, people have often assumed that there is an abstract feature shared by structural cases such as accusative and nominative and one which is shared by obligue cases such as dative or accusative.
- And in order to do justice to the observation that dative and accusative share some kind of feature, they assumed that accusative and dative share a feature to the exclusion of nominative and genitive.

(15)
a. Nominative = [–oblique,–object]
b. Accusative = [–oblique,+object]
c. Dative = [+oblique,+object]
d. Genitive = [+oblique,–object]

(16)
a. Nominative = [–oblique,–object]
b. Accusative = [–oblique,+object]
c. Dative = [+oblique,+object]
d. Genitive = [+oblique,–object]

- In such a system, we can refer to the two cases accusative and nominative with the feature [–oblique]

(16)
a. Nominative = [–oblique,–object]
b. Accusative = [–oblique,+object]
c. Dative = [+oblique,+object]
d. Genitive = [+oblique,–object]

- In such a system, we can refer to the two cases accusative and nominative with the feature [–oblique]
- These two cases form a *natural class* with respect to the feature.

- Coming back to the paradigm of German determiner inflection, we can thus assume that the determiner /die/ could be specified as follows:

(17) /die/ $\Leftrightarrow$ [fem, –obl]

(18)

	Masc	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- The same can be done for the form /das/:

(19) /das/ $\Leftrightarrow$ [neut, –obl]

(20)

	Mask	Fem	Neut
Nom	der	die	das
Akk	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- The same can be done for the form /das/:

(19) /das/ ⇔ [neut, –obl]

(20)

	Mask	Fem	Neut
Nom	der	die	das
Akk	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- And the form /der/:

(21) /der/ ⇔ [fem, +obl]

(22)

	Mask	Fem	Neut
Nom	der	die	das
Akk	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- We are not entirely satisfied yet. The data suggest that we also need a possibility to refer to masculine and neuter to the exclusion of feminine.

(23)

	Mask	Fem	Neut
Nom	der	die	das
Akk	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- We are not entirely satisfied yet. The data suggest that we also need a possibility to refer to masculine and neuter to the exclusion of feminine.

(23)

	Mask	Fem	Neut
Nom	der	die	das
Akk	den	die	das
Dat	dem	der	dem
Gen	des	der	des

- This suggests that we also need to decompose the feature gender.

- For lack of other labels, people have used the labels [±fem] and [±masc]:

(24)
a. Masculine = [+masc,–fem]
b. Feminine = [–masc,+fem]
c. Neuter = [–masc,–fem]

- For lack of other labels, people have used the labels [±fem] and [±masc]:

(24)
a. Masculine = [+masc,–fem]
b. Feminine = [–masc,+fem]
c. Neuter = [–masc,–fem]

- This allows us to refer to the set of masculine and neuter with the feature [–fem].

- So let us go through the whole paradigm of German determiner inflection:

(25)

	Mask	Fem	Neut
Nom	der	die	das
Akk	den	die	das
Dat	dem	der	dem
Gen	des	der	des

(26) Case:

Nom:	[–obj]	[–obl]
Acc:	[+obj]	[–obl]
Dat:	[+obj]	[+obl]
Gen:	[–obj]	[+obl]

(27) Gender:

M:	[+masc]	[–fem]
F:	[–masc]	[+fem]
N:	[–masc]	[–fem]

- We begin with the fully specified Vocabulary Items:

	Masc	Fem	Neut
Nom	der		
Acc	den		
Dat			
Gen			

Vocabulary Items:

a. /der/⇔ [–obj,–obl,+masc,–fem]
b. /den/⇔ [+obj,–obl,+masc,–fem]

- Now for the underspecified markers in the structural (non-oblique) cases.

	Masc	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat			
Gen			

Vocabulary Items:

a. /der/⇔ [–obj,–obl,+masc,–fem]
b. /den/⇔ [+obj,–obl,+masc,–fem]
c. /die/⇔ [–obl,–masc,+fem]
d. /das/⇔ [–obl,–masc,–fem]

- Now for the underspecified markers in the oblique cases:

	Masc	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat		der	
Gen		der	

Vocabulary Items:

a. /der/⇔ [–obj,–obl,+masc,–fem]
b. /den/⇔ [+obj,–obl,+masc,–fem]
c. /die/⇔ [–obl,–masc,+fem]
d. /das/⇔ [–obl,–masc,–fem]
e. /der/⇔ [+obl,–masc,+fem]

- Now the exponents underspecified for gender:

	Masc	Fem	Neut
Nom	der	die	das
Acc	den	die	das
Dat	dem	der	dem
Gen	des	der	des

Vocabulary Items:

a. /der/⇔ [–obj,–obl,+masc,–fem]
b. /den/⇔ [+obj,–obl,+masc,–fem]
c. /die/⇔ [–obl,–masc,+fem]
d. /das/⇔ [–obl,–masc,–fem]
e. /der/⇔ [+obl,–masc,+fem]
f. /dem/⇔ [+obl,+obj,–fem]
g. /des/⇔ [+obl,–obj,–fem]

- We managed to model the distribution of markers using 7 Vocabulary Items.

- We managed to model the distribution of markers using 7 Vocabulary Items.
- If we didn't assume Underspecification, the Subset Principle or decomposition, but rather assumed that each and every cell would be stored on its own, then we would have to assume 12 markers for the same paradigm.
- Our model is clearly more economical.

- That becomes even more apparent when we consider the plural paradigm, which completely neutralizes all gender distinctions.

(28)

		Masc	Fem	Neut
SG	Nom	der	die	das
	Acc	den	die	das
	Dat	dem	der	dem
	Gen	des	der	des
PL	Nom	die	die	die
	Acc	die	die	die
	Dat	den	den	den
	Gen	der	der	der

- That becomes even more apparent when we consider the plural paradigm, which completely neutralizes all gender distinctions.

(28)

		Masc	Fem	Neut
SG	Nom	der	die	das
	Acc	den	die	das
	Dat	dem	der	dem
	Gen	des	der	des
PL	Nom	die	die	die
	Acc	die	die	die
	Dat	den	den	den
	Gen	der	der	der

- It is very unlikely that speakers of German store each and every feature combination independently. That would be very redundant.

2. Case Study: Inflection of Icelandic Nouns:

(29) Icelandic Noun Inflection:

	Masc	Fem	Neut
Nom.SG	-ur	-∅	-∅
Acc.SG	-∅	-∅	-∅
Dat.SG	-i	-∅	-i
Gen.SG	-s	-ar	-s
Nom.PL	-ar	-ir	-∅
Acc.PL	-a	-ir	–∅
Dat.PL	-um	-um	-um
Gen.PL	-a	-a	-a

- We can take over the decomposition of gender and case from German.

Icelandic Noun Inflection:

	Masc	Fem	Neut
Nom.SG	-ur	-∅	-∅
Acc.SG	-∅	-∅	-∅
Dat.SG	-i	-∅	-i
Gen.SG	-s	-ar	-s
Nom.PL	-ar	-ir	-∅
Acc.PL	-a	-ir	-∅
Dat.PL	-um	-um	-um
Gen.PL	-a	-a	-a

VIs:

a. /-ur/⇔ [–obj,–obl,+masc,–fem,sg]
b. /-ar/⇔ [–obj,+obl,–masc,+fem,sg]
c. /-a/⇔ [+obj,–obl,+masc,–fem,pl]
d. /-ar/⇔ [–obj,–obl,+masc,–fem,pl]
e. /-ir/⇔ [–obj,–obl,–masc,+fem,pl]
f. /-um/⇔ [+obj,+obl,pl]
g. /-a/⇔ [–obj,+obl,pl]
h. /-i/⇔ [+obj,+obl,–fem]
i. /-s/⇔ [–obj,+obl,–fem]
j. /∅/ ⇔ [ ]

- Again, we only need 10 VIs to fill the paradigm. If we didn't assume underspecification, etc. we would need to story 24 lexical entries.
- And we can straightforwardly see the use of all of the concepts we saw today:
  - ↪ We can see that the distribution of some markers (e.g. genitive or dative plural) can be captured with underspecification

- Again, we only need 10 VIs to fill the paradigm. If we didn't assume underspecification, etc. we would need to story 24 lexical entries.
- And we can straightforwardly see the use of all of the concepts we saw today:
  - ↪ We can see that the distribution of some markers (e.g. genitive or dative plural) can be captured with underspecification
  - ↪ The decomposition of case allows us to capture the systematic syncretism of structural cases

- Again, we only need 10 VIs to fill the paradigm. If we didn't assume underspecification, etc. we would need to story 24 lexical entries.
- And we can straightforwardly see the use of all of the concepts we saw today:
  - ↪ We can see that the distribution of some markers (e.g. genitive or dative plural) can be captured with underspecification
  - ↪ The decomposition of case allows us to capture the systematic syncretism of structural cases
  - ↪ The decomposition of gender allows us to model the syncretism of masculine and neuter to the exclusion of feminine (dative singular)

- Again, we only need 10 VIs to fill the paradigm. If we didn't assume underspecification, etc. we would need to story 24 lexical entries.
- And we can straightforwardly see the use of all of the concepts we saw today:
  - ↪ We can see that the distribution of some markers (e.g. genitive or dative plural) can be captured with underspecification
  - ↪ The decomposition of case allows us to capture the systematic syncretism of structural cases
  - ↪ The decomposition of gender allows us to model the syncretism of masculine and neuter to the exclusion of feminine (dative singular)
  - ↪ The use of the Subset Principle is most obvious with the elsewhere marker /∅/ which occurs in all kinds of places.

Side Remarks:

- The fact that we used the same decomposition for Icelandic and German does not follow from anything per se.
  - ↪ Sometimes it is assumed that we want to assume a universal decomposition of some morphosyntactic categories.
  - ↪ However, that assumption often proves problematic since other languages have other patterns of syncretism.

Side Remarks:

- The fact that we used the same decomposition for Icelandic and German does not follow from anything per se.
  - ↪ Sometimes it is assumed that we want to assume a universal decomposition of some morphosyntactic categories.
  - ↪ However, that assumption often proves problematic since other languages have other patterns of syncretism.
- First and foremost, the decomposition of morphosyntactic categories should be motivated empirically.

Side Remarks:

- Also, in Icelandic we had to posit two homophonous markers /a/. The same was true in German where we assumed two instances of /der/.
  - ↪ Such cases are often referred to as *unresolved syncretism*.
  - ↪ These cases of syncretism can (presumably) not be modelled with the current tools we have developed so far.
  - ↪ The assumption is that some instances of syncretism are accidental.
- Often, it is not clear which cases of syncretism count as systematic and which count as accidental.

Side Remarks:

- Also, in Icelandic we had to posit two homophonous markers /a/. The same was true in German where we assumed two instances of /der/.
  - ↪ Such cases are often referred to as *unresolved syncretism*.
  - ↪ These cases of syncretism can (presumably) not be modelled with the current tools we have developed so far.
  - ↪ The assumption is that some instances of syncretism are accidental.
- Often, it is not clear which cases of syncretism count as systematic and which count as accidental.
- In practice, we often let the theory decide and assume accidental homophony only in cases where our theories are not powerful enough to derive them.

Side Remarks:

- In Icelandic, we found that the radically underspecified morphological marker was also phonologically radically unmarked (i.e. /∅/).
  - ↪ Again, nothing in our theory so far says that this is necessary.
- However, it has been claimed that an elegant analysis relates morphological complexity with phonological complexity (The Iconicity Principle in Wiese (1996,1999)).

- But again, such assumptions are crosslinguistically often problematic since we find several cases where we would rather have a different marker radically underspecified.

Singular:

	I	II	III
nom	-∅	-a	-∅
gen	-a	-ē	-i
dat	-u	-i	-i
acc	≃nom	-u	≃nom
loc	-u	-i	-i
instr	-om/em	-ōm	-ju/-i
voc	-e/-u	-o	-i

Plural:

	I	II	III
nom	-i	-e	-i
gen	-ā	-ā	-ī
dat	-i-ma	-i-ma	-i-ma
acc	e	-e	-i
loc	-i-ma	-a-ma	-i-ma
instr	-i-ma	-a-ma	-i-ma
voc	-i	-e	-i

- If we take the Subset Principle seriously, we can erase quite a number of features from the specifications:

(30)

	A	B
$-\alpha$	♠	♠
$+\alpha$	♦	♦

- In principle, it would suffice to fully specify one of the markers in (30) and have the other one be the elsewhere marker.

(31) a. $[-\alpha] \leftrightarrow$ /♠/
b. [ ] $\leftrightarrow$ /♦/

- The Subset Principle will do the rest

- The same holds for our case study from Icelandic:

Maximally specified VIs:

a. /-ur/⇔ [–obj,–obl,+masc,–fem,sg]
b. /-ar/⇔ [–obj,+obl,–masc,+fem,sg]
c. /-a/⇔ [+obj,–obl,+masc,–fem,sg]
d. /-ar/⇔ [–obj,–obl,+masc,–fem,pl]
e. /-ir/⇔ [–obj,–obl,–masc,+fem,pl]
f. /-i/⇔ [+obj,+obl,–fem,sg]
g. /-s/⇔ [–obj,+obl,–fem,sg]
h. /-um/⇔ [+obj,+obl,pl]
i. /-a/⇔ [–obj,+obl,pl]
j. /∅/ ⇔ [ ]

Minimally specified VIs:

a. /-ur/⇔ [–obj,–obl,+masc,sg]
b. /-ar/⇔ [–obj,+obl,+fem,sg]
c. /-a/⇔ [+obj,–obl,+masc,sg]
d. /-ar/⇔ [–obj,–obl,+masc]
e. /-ir/⇔ [–obj,–obl,+fem,pl]
f. /-i/⇔ [+obj,+obl,–fem,sg]
g. /-s/⇔ [+obl,–fem,sg]
h. /-um/⇔ [+obj,+obl]
i. /-a/⇔ [+obl,pl]
j. /∅/ ⇔ [ ]

- Empirically, these two lists of VIs make the exact same predictions. Both result in the same distribution of markers.
- Halle 1997 states that markers should be minimally specified:

(32) The number of features mentioned in the Vocabulary must be minimized.

Halle 1997:130

- That, however, is merely a stipulation.
- A lot about this decision depends on one's theory of learning:
  - ↪ Do children learn VIs by postulating maximally specified VIs and then deleting features when confronted with syncretism?
  - ↪ Or do children learn VIs by postulating syncretism throughout and only posit additional VIs when confronted with discriminating forms?

- That, however, is merely a stipulation.
- A lot about this decision depends on one's theory of learning:
  - ↪ Do children learn VIs by postulating maximally specified VIs and then deleting features when confronted with syncretism?
  - ↪ Or do children learn VIs by postulating syncretism throughout and only posit additional VIs when confronted with discriminating forms?
- See dissertation by Pertsova 2007

The case studies above required a decomposition of gender and case.

- But other morphosyntactic categories are often decomposed as well
- Person is often decomposed into [±speaker] and [±participant].

(33) A decomposition of person:

a. [+speaker,+participant] = 1.Person inclusive
b. [+speaker,−participant] = 1.Person exclusive
c. [−speaker,+participant] = 2.Person
d. [−speaker,−participant] = 3.Person

Another possible decomposition uses the features [±1] and [±2].

- The decomposition would then look like this:

(34) Decomposition into [±1] and [±2]:

a. [+1],[−2] = 1. Person
b. [+1],[+2] = 1. Person inclusive
c. [−1],[+2] = 2. Person
d. [−1],[−2] = 3. Person

- Unlike with the person decomposition in (33), (34) allows us to capture the germanic 1+3 syncretism.

(35)

	SG	PL
1	was	were
2	were	were
3	was	were

It should be noted though that this 1+3-syncretism is very rare outside of Germanic (see Corbett 2006).

- Much more common is a syncretism of 1 and 2 as opposed to 3.

(36) Nez Perce Verbal Agreement:

'go'	SG	PL
1	kiyú?	pe-kiyú?
2	kiyú?	pe-kiyú?
3	hi-kiyú?	hi-pe-kiyú?

- For this reason, many people prefer the decomposition into [±speaker] und [±participant] which allows us to capture the pattern in (36).

Most systems of number do not require a decomposition since there are typically only two values (SG and PL). Thus a decomposition does not help resolve syncretism.

- But as soon as more values are involved, it can become helpful.
- Systems with a dual are sometimes decomposed into [±Singular] and [±Plural]:

(37) Eine Dekomposition von Numerus:

a. [+Singular,–Plural] = Singular
b. [–Singular,–Plural] = Dual
c. [–Singular,+Plural] = Plural

- The forth logical combination is contradictory.

- In some languages, it has been claimed that we find even more complex systems such as a trial or a paucal.
- Those systems cannot be captured with [±SG] with [±PL].

- In some languages, it has been claimed that we find even more complex systems such as a trial or a paucal.
- Those systems cannot be captured with [±SG] with [±PL].
- More recent works adopt a decomposition into [±Singular] and [±Augmented] for simple systems or [±Singular],[±Plural] and [±Augmented] for more complex ones (see Harbour 2006,2011 for discussion).

- In some languages, it has been claimed that we find even more complex systems such as a trial or a paucal.
- Those systems cannot be captured with [±SG] with [±PL].
- More recent works adopt a decomposition into [±Singular] and [±Augmented] for simple systems or [±Singular],[±Plural] and [±Augmented] for more complex ones (see Harbour 2006,2011 for discussion).
- It should be noted though that it also has been claimed that trial or paucal are not actual linguistic categories.

- Verbal categories such as aspect, tense, modality, etc. are also sometimes decomposed but too few case studies have been proposed.

- Verbal categories such as aspect, tense, modality, etc. are also sometimes decomposed but too few case studies have been proposed.
  - ↪ But the crosslinguistic variation (and their highly flexible morphosemantics) across tense or aspect systems often makes it hard to come up with crosslinguistically applicable feature systems.

- Verbal categories such as aspect, tense, modality, etc. are also sometimes decomposed but too few case studies have been proposed.
  - ↪ But the crosslinguistic variation (and their highly flexible morphosemantics) across tense or aspect systems often makes it hard to come up with crosslinguistically applicable feature systems.
- Alexiadou & Müller (2005) also propose a decomposition for inflection class features in Greek and Russian.
  - ↪ But since inflection class is typically viewed as an arbitrary lexical category (that often interacts with phonology to a certain degree) many people think that these features are not actually part of the morphosyntax.

Conclusion:

- We saw several tools to capture syncretism patterns:

Conclusion:

- We saw several tools to capture syncretism patterns:
  - Underspecification
  - The Subset Principle
  - Decomposition

Conclusion:

- We saw several tools to capture syncretism patterns:
  - Underspecification
  - The Subset Principle
  - Decomposition
- In combination, these tools provide for a powerful (but not unconstrained) system to capture various syncretism patterns.