

Item and Pattern Morphology

James P. Blevins

University of Cambridge
jpb39@cam.ac.uk

Quantitative Measures in Morphology & Morphological
Development, 15.01.11

Outline

- 1 Morphological analysis
 - Types of phenomena
 - Conceptions of analysis
- 2 Models of grammatical analysis
 - Morphemic models
 - Item and pattern models
 - Stem-based implication
- 3 Implicational analysis
 - System factorization
 - Fractured noun declensions in German
 - Modelling implicational structure
- 4 Summary and implications

Items and patterns

- Why are item/pattern models relevant to quantitative measures?
 - Because they provide appropriate items to count, i.e., words.
- Why are these models relevant to morphological development?
 - Because they incorporate a speaker-oriented perspective, which addresses tasks that speakers face in language acquisition and use.

Items and patterns

- Why are item/pattern models relevant to quantitative measures?
 - Because they provide appropriate items to count, i.e., words.
- Why are these models relevant to morphological development?
 - Because they incorporate a speaker-oriented perspective, which addresses tasks that speakers face in language acquisition and use.

Items and patterns

- Why are item/pattern models relevant to quantitative measures?
 - Because they provide appropriate items to count, i.e., words.
- Why are these models relevant to morphological development?
 - Because they incorporate a speaker-oriented perspective, which addresses tasks that speakers face in language acquisition and use.

Items and patterns

- Why are item/pattern models relevant to quantitative measures?
 - Because they provide appropriate items to count, i.e., words.
- Why are these models relevant to morphological development?
 - Because they incorporate a speaker-oriented perspective, which addresses tasks that speakers face in language acquisition and use.

Speaker-oriented analysis

- Morphological analysis models the morphological information that speakers exploit in the acquisition and use of a language.
- Speakers need to be able to produce and interpret the forms of a language, including forms that they may not have encountered.
- In languages with rich inflectional morphology, this is the ‘paradigm cell filling problem’ (Ackerman et al. 2009).
- The difficulty of this task does not appear to differ significantly across languages (Malouf & Ackerman 2010a,b), presumably reflecting adaptive pressures on language acquisition and use.

Speaker-oriented analysis

- Morphological analysis models the morphological information that speakers exploit in the acquisition and use of a language.
- Speakers need to be able to produce and interpret the forms of a language, including forms that they may not have encountered.
- In languages with rich inflectional morphology, this is the 'paradigm cell filling problem' (Ackerman et al. 2009).
- The difficulty of this task does not appear to differ significantly across languages (Malouf & Ackerman 2010a,b), presumably reflecting adaptive pressures on language acquisition and use.

Speaker-oriented analysis

- Morphological analysis models the morphological information that speakers exploit in the acquisition and use of a language.
- Speakers need to be able to produce and interpret the forms of a language, including forms that they may not have encountered.
- In languages with rich inflectional morphology, this is the 'paradigm cell filling problem' (Ackerman et al. 2009).
- The difficulty of this task does not appear to differ significantly across languages (Malouf & Ackerman 2010a,b), presumably reflecting adaptive pressures on language acquisition and use.

Speaker-oriented analysis

- Morphological analysis models the morphological information that speakers exploit in the acquisition and use of a language.
- Speakers need to be able to produce and interpret the forms of a language, including forms that they may not have encountered.
- In languages with rich inflectional morphology, this is the 'paradigm cell filling problem' (Ackerman et al. 2009).
- The difficulty of this task does not appear to differ significantly across languages (Malouf & Ackerman 2010a,b), presumably reflecting adaptive pressures on language acquisition and use.

Theoretical lexicography

- Many other familiar properties of morphological descriptions fall under what might be called ‘theoretical lexicography’:
 - Speakers do NOT need to be able to associate a single meaning with individual formatives, to assign invariant segmentations to forms, to identify a unique set of principal parts, to assign forms to a fixed number of inflection classes, etc.
- These properties of languages exhibit far greater variation, presumably reflecting the lack of adaptive pressures.

Theoretical lexicography

- Many other familiar properties of morphological descriptions fall under what might be called 'theoretical lexicography':
 - Speakers do NOT need to be able to associate a single meaning with individual formatives, to assign invariant segmentations to forms, to identify a unique set of principal parts, to assign forms to a fixed number of inflection classes, etc.
- These properties of languages exhibit far greater variation, presumably reflecting the lack of adaptive pressures.

Theoretical lexicography

- Many other familiar properties of morphological descriptions fall under what might be called 'theoretical lexicography':
 - Speakers do NOT need to be able to associate a single meaning with individual formatives, to assign invariant segmentations to forms, to identify a unique set of principal parts, to assign forms to a fixed number of inflection classes, etc.
- These properties of languages exhibit far greater variation, presumably reflecting the lack of adaptive pressures.

Atomistic (Neo-Bloomfieldian) approaches

- Language acquisition involves learning form-meaning mappings.
- Language use involves the interpretation of forms and the realization of meanings in isolation from other forms/meanings.
- The lexicon is a redundancy-free repository of meaningful units.

Atomistic (Neo-Bloomfieldian) approaches

- Language acquisition involves learning form-meaning mappings.
- Language use involves the interpretation of forms and the realization of meanings in isolation from other forms/meanings.
- The lexicon is a redundancy-free repository of meaningful units.

Atomistic (Neo-Bloomfieldian) approaches

- Language acquisition involves learning form-meaning mappings.
- Language use involves the interpretation of forms and the realization of meanings in isolation from other forms/meanings.
- The lexicon is a redundancy-free repository of meaningful units.

Neo-Neogrammarian models

- Language acquisition involves learning (i) form inventories, the (ii) distribution of forms in paradigms and in syntactic contexts and (iii) the functions and meanings associated with forms.
- Language use involves highly aggressive prediction (possibly to minimize/reduce the amount of real-time processing of input).
- The lexicon is a statistical inferencing engine that provides an analogical base for extending encountered patterns to new items.

Neo-Neogrammarian models

- Language acquisition involves learning (i) form inventories, the (ii) distribution of forms in paradigms and in syntactic contexts and (iii) the functions and meanings associated with forms.
- Language use involves highly aggressive prediction (possibly to minimize/reduce the amount of real-time processing of input).
- The lexicon is a statistical inferencing engine that provides an analogical base for extending encountered patterns to new items.

Neo-Neogrammarian models

- Language acquisition involves learning (i) form inventories, the (ii) distribution of forms in paradigms and in syntactic contexts and (iii) the functions and meanings associated with forms.
- Language use involves highly aggressive prediction (possibly to minimize/reduce the amount of real-time processing of input).
- The lexicon is a statistical inferencing engine that provides an analogical base for extending encountered patterns to new items.

Implicational morphology

- Implicational approaches model speaker-oriented analysis.
- ‘Word and Paradigm’ models are pure implicational approaches.
- Words and paradigms merely sanction reliable implications.

Implicational morphology

- Implicational approaches model speaker-oriented analysis.
- 'Word and Paradigm' models are pure implicational approaches.
- Words and paradigms merely sanction reliable implications.

Implicational morphology

- Implicational approaches model speaker-oriented analysis.
- 'Word and Paradigm' models are pure implicational approaches.
- Words and paradigms merely sanction reliable implications.

Item and arrangement

The essence of IA is to talk simply of things and the arrangements in which those things occur (Hockett 1954: 387)

The grammar, or grammatical system, of a language is (1) THE MORPHEMES USED IN THE LANGUAGE, and (2) THE ARRANGEMENTS IN WHICH THESE MORPHEMES OCCUR RELATIVE TO EACH OTHER IN UTTERANCES. (Hockett 1958: 129)

Item and arrangement

The essence of IA is to talk simply of things and the arrangements in which those things occur (Hockett 1954: 387)

The grammar, or grammatical system, of a language is (1) THE MORPHEMES USED IN THE LANGUAGE, and (2) THE ARRANGEMENTS IN WHICH THESE MORPHEMES OCCUR RELATIVE TO EACH OTHER IN UTTERANCES. (Hockett 1958: 129)

Problems of morphemic analysis

One motive for the post-Bloomfieldian model consisted, that is to say, in a genuinely factual assertion about language: namely, that there is some sort of matching between minimal 'sames' of 'form' (morphs) and 'meaning' (morphemes). Qua factual assertion this has subsequently proved false: for certain languages, such as Latin, the correspondence which was envisaged apparently does not exist ... One is bound to suspect, in the light of such a conclusion, that the model is in some sense wrong. (Matthews 1972: 124)

Item and process

Morphophonemes, morphs, phones, and acoustic phones are ARTIFACTS OF ANALYSIS or CONVENIENCES FOR DESCRIPTION, not elements in a language. (Hockett 1961: 42)

Some of the phonemic material in a derived form may be, not part of any underlying form, but rather a REPRESENTATION or MARKER of the process. (Hockett 1954: 396)

Item and process

Morphophonemes, morphs, phones, and acoustic phones are ARTIFACTS OF ANALYSIS or CONVENIENCES FOR DESCRIPTION, not elements in a language. (Hockett 1961: 42)

Some of the phonemic material in a derived form may be, not part of any underlying form, but rather a REPRESENTATION or MARKER of the process. (Hockett 1954: 396)

'WP' models

Quite apart from minor variants of IP or IA, or models that might be invented tomorrow, there is one model which is clearly distinct from either IA or IP, and which is older and more respectable than either. This is the WORD AND PARADIGM (WP) model, the traditional framework for the discussion of Latin, Greek, Sanskrit, and a good many more modern familiar languages. ... (Hockett 1954: 386)

- An inflectional system is factored into a set of exemplary patterns and a set of diagnostic principal parts for non-exemplary items.

'WP' models

Quite apart from minor variants of IP or IA, or models that might be invented tomorrow, there is one model which is clearly distinct from either IA or IP, and which is older and more respectable than either. This is the WORD AND PARADIGM (WP) model, the traditional framework for the discussion of Latin, Greek, Sanskrit, and a good many more modern familiar languages. ... (Hockett 1954: 386)

- An inflectional system is factored into a set of exemplary patterns and a set of diagnostic principal parts for non-exemplary items.

A (slight) mischaracterization

- 'Item and arrangement' and 'item and process' both refer to units ('items') and modes of combination ('arrangements'/'processes').
- But 'word and paradigm' refers to a PARTICULAR unit ('word') and a SPECIFIC network of units (inflectional 'paradigm').
- The third model should be 'item and pattern', where comparison of the item against the pattern sanctions the deduction of forms.

A (slight) mischaracterization

- 'Item and arrangement' and 'item and process' both refer to units ('items') and modes of combination ('arrangements'/'processes').
- But 'word and paradigm' refers to a PARTICULAR unit ('word') and a SPECIFIC network of units (inflectional 'paradigm').
- The third model should be 'item and pattern', where comparison of the item against the pattern sanctions the deduction of forms.

A (slight) mischaracterization

- 'Item and arrangement' and 'item and process' both refer to units ('items') and modes of combination ('arrangements'/'processes').
- But 'word and paradigm' refers to a PARTICULAR unit ('word') and a SPECIFIC network of units (inflectional 'paradigm').
- The third model should be 'item and pattern', where comparison of the item against the pattern sanctions the deduction of forms.

Why words?

- Words are often of greater predictive value than sub-word units.
 - They are IDENTIFIABLE because they are more consistently demarcated in the speech stream than sub-word units.
 - They are more unambiguously INTERPRETABLE than sub-word units.
 - The interpretation of a morphological unit is just not the sum of the interpretation of its parts: COMBINATIONS may be distinctive.

Why words?

- Words are often of greater predictive value than sub-word units.
 - They are IDENTIFIABLE because they are more consistently demarcated in the speech stream than sub-word units.
 - They are more unambiguously INTERPRETABLE than sub-word units.
 - The interpretation of a morphological unit is just not the sum of the interpretation of its parts: COMBINATIONS may be distinctive.

Why words?

- Words are often of greater predictive value than sub-word units.
 - They are IDENTIFIABLE because they are more consistently demarcated in the speech stream than sub-word units.
 - They are more unambiguously INTERPRETABLE than sub-word units.
- The interpretation of a morphological unit is just not the sum of the interpretation of its parts: COMBINATIONS may be distinctive.

Why words?

- Words are often of greater predictive value than sub-word units.
 - They are IDENTIFIABLE because they are more consistently demarcated in the speech stream than sub-word units.
 - They are more unambiguously INTERPRETABLE than sub-word units.
 - The interpretation of a morphological unit is just not the sum of the interpretation of its parts: COMBINATIONS may be distinctive.

Why words?

The word is a more stable and solid focus of grammatical relations than the component morpheme by itself. Put another way, grammatical statements are abstractions, but they are more profitably abstracted from words as wholes than from individual morphemes. (Robins 1959: 128).

Why paradigms?

- Paradigms exhibit the most reliable patterns of interpredictability because they are defined over a closed, uniform feature space:
 - Notions like ‘morphological gap’, ‘suppletion’, and even ‘syncretism’ are mainly applied to inflectional paradigms.
 - Conversely, inflected forms are rarely taken to be ‘established’.

Why paradigms?

- Paradigms exhibit the most reliable patterns of interpredictability because they are defined over a closed, uniform feature space:
 - Notions like ‘morphological gap’, ‘suppletion’, and even ‘syncretism’ are mainly applied to inflectional paradigms.
 - Conversely, inflected forms are rarely taken to be ‘established’.

Why paradigms?

- Paradigms exhibit the most reliable patterns of interpredictability because they are defined over a closed, uniform feature space:
 - Notions like ‘morphological gap’, ‘suppletion’, and even ‘syncretism’ are mainly applied to inflectional paradigms.
 - Conversely, inflected forms are rarely taken to be ‘established’.

The scope of implication

- Yet ‘word and paradigm’ is a specific instance of an ‘item and pattern’ model and debates about the morphological status of words and paradigms, though important, are subsidiary:
 - In periphrastic constructions (Ackerman & Stump 2004), units larger than the word may be an equally “stable and solid focus of grammatical relations” (cf. Robins 1959).
 - A perfectly agglutinative system would also sanction reliable predictions between ‘units of meaning’ and ‘units of form’.
 - Derivational paradigms and other types of morphological families also exhibit (a typically weaker) implicational structure.

The scope of implication

- Yet ‘word and paradigm’ is a specific instance of an ‘item and pattern’ model and debates about the morphological status of words and paradigms, though important, are subsidiary:
 - In periphrastic constructions (Ackerman & Stump 2004), units larger than the word may be an equally “stable and solid focus of grammatical relations” (cf. Robins 1959).
 - A perfectly agglutinative system would also sanction reliable predictions between ‘units of meaning’ and ‘units of form’.
 - Derivational paradigms and other types of morphological families also exhibit (a typically weaker) implicational structure.

The scope of implication

- Yet ‘word and paradigm’ is a specific instance of an ‘item and pattern’ model and debates about the morphological status of words and paradigms, though important, are subsidiary:
 - In periphrastic constructions (Ackerman & Stump 2004), units larger than the word may be an equally “stable and solid focus of grammatical relations” (cf. Robins 1959).
 - A perfectly agglutinative system would also sanction reliable predictions between ‘units of meaning’ and ‘units of form’.
 - Derivational paradigms and other types of morphological families also exhibit (a typically weaker) implicational structure.

The scope of implication

- Yet ‘word and paradigm’ is a specific instance of an ‘item and pattern’ model and debates about the morphological status of words and paradigms, though important, are subsidiary:
 - In periphrastic constructions (Ackerman & Stump 2004), units larger than the word may be an equally “stable and solid focus of grammatical relations” (cf. Robins 1959).
 - A perfectly agglutinative system would also sanction reliable predictions between ‘units of meaning’ and ‘units of form’.
 - Derivational paradigms and other types of morphological families also exhibit (a typically weaker) implicational structure.

Stem and pattern systems?

- Sub-word units consisting of roots and formatives are often informative in ways that roots in isolation are not, leading to analyses based on abstract stems (Aronoff 1994), stem sets (Anderson 1992) or stem spaces (Bonami & Boyé 2003).
- Yet stem-based approaches create a pair of new challenges:
 - The demarcation of stems raises new difficulties of segmentation.
 - The interpretation and predictive value of stems often depends on their distribution within full word forms.

Stem and pattern systems?

- Sub-word units consisting of roots and formatives are often informative in ways that roots in isolation are not, leading to analyses based on abstract stems (Aronoff 1994), stem sets (Anderson 1992) or stem spaces (Bonami & Boyé 2003).
- Yet stem-based approaches create a pair of new challenges:
 - The demarcation of stems raises new difficulties of segmentation.
 - The interpretation and predictive value of stems often depends on their distribution within full word forms.

Stem and pattern systems?

- Sub-word units consisting of roots and formatives are often informative in ways that roots in isolation are not, leading to analyses based on abstract stems (Aronoff 1994), stem sets (Anderson 1992) or stem spaces (Bonami & Boyé 2003).
- Yet stem-based approaches create a pair of new challenges:
 - The demarcation of stems raises new difficulties of segmentation.
 - The interpretation and predictive value of stems often depends on their distribution within full word forms.

Stem and pattern systems?

- Sub-word units consisting of roots and formatives are often informative in ways that roots in isolation are not, leading to analyses based on abstract stems (Aronoff 1994), stem sets (Anderson 1992) or stem spaces (Bonami & Boyé 2003).
- Yet stem-based approaches create a pair of new challenges:
 - The demarcation of stems raises new difficulties of segmentation.
 - The interpretation and predictive value of stems often depends on their distribution within full word forms.

Problems of segmentation I

- The ‘recalcitrance’ of English *children*:
 - 1 *child + ren* ~ *childr + en* — “each of the points of division has advantages and disadvantages” (Harris 1942 113).
 - 2 *child + r + en* ~ *child + ablaut + en* ~ *children* — “this is one of the cases in which all of our preferential criteria ... fail and nothing remains but a resort to convenience” (Hockett 1947: 240).

Problems of segmentation I

- The ‘recalcitrance’ of English *children*:
 - ① *child + ren* ~ *childr + en* — “each of the points of division has advantages and disadvantages” (Harris 1942 113).
 - ② *child + r + en* ~ *child + ablaut + en* ~ *children* — “this is one of the cases in which all of our preferential criteria ... fail and nothing remains but a resort to convenience” (Hockett 1947: 240).

Problems of segmentation I

- The ‘recalcitrance’ of English *children*:
 - 1 *child* + *ren* ~ *childr* + *en* — “each of the points of division has advantages and disadvantages” (Harris 1942 113).
 - 2 *child* + *r* + *en* ~ *child* + ablaut + *en* ~ *children* — “this is one of the cases in which all of our preferential criteria ... fail and nothing remains but a resort to convenience” (Hockett 1947: 240).

Problems of segmentation II

The order of morphemes is fixed [in Spanish]: (derivational prefix(es)) + lexical stem + theme vowel + tense marker (sometimes including an empty morph) + person marker. Some forms, however, have fused in the course of history and a neat segmentation is not always possible. The preterit is the most difficult paradigm to analyse, since the theme vowel is sometimes indistinguishable, and segmenting the second and third person plural markers in the regular way, /-is, -n/, leaves an awkward residue that occurs nowhere else in the system. (Green 1997: 99)

Parasitism

	Sg	Pl	Sg	Pl
Nom	kuotam	kuotamazh	kuorta	kuortozh
Gen	kuotama	kuotamii	kerta	kuortoi
Dat	kuotamaa	kuotamazhta	kertaa	kuortozhta
Erg	kuotamuo	kuotamazh	kertuo	kuortozh
All	kuotamaga	kuotamazhka	kertaga	kuortozhka
Ins	kuotamaca	kuotamazhca	kertaca	kuortozhca
Lat	kuotamagh	kuotamegh	kertagh	kuortuojegh
Ins	kuotamal	kuotamel	kertal	kuortuojel
	hen (1)		head (3)	

Ingush noun declensions (Nichols 2011)

Predictive syncretism in Ingush

- In the singular, a genitive form X predicts (and is predicted by):
 - a dative $X + a$, an allative $X + ga$ and an instrumental $X + ca$.
- In the plural, an ergative form X predicts (and is predicted by):
 - a dative $X + ta$, an allative $X + ka$ and an instrumental $X + ca$.

Predictive syncretism in Ingush

- In the singular, a genitive form X predicts (and is predicted by):
 - a dative $X + a$, an allative $X + ga$ and an instrumental $X + ca$.
- In the plural, an ergative form X predicts (and is predicted by):
 - a dative $X + ta$, an allative $X + ka$ and an instrumental $X + ca$.

Predictive syncretism in Ingush

- In the singular, a genitive form X predicts (and is predicted by):
 - a dative $X + a$, an allative $X + ga$ and an instrumental $X + ca$.
- In the plural, an ergative form X predicts (and is predicted by):
 - a dative $X + ta$, an allative $X + ka$ and an instrumental $X + ca$.

Predictive syncretism in Ingush

- In the singular, a genitive form X predicts (and is predicted by):
 - a dative $X + a$, an allative $X + ga$ and an instrumental $X + ca$.
- In the plural, an ergative form X predicts (and is predicted by):
 - a dative $X + ta$, an allative $X + ka$ and an instrumental $X + ca$.

Parasitism

	Sg	Pl	Sg	Pl
Nom	kuotam	kuotamazh	kuorta	kuortozh
Gen	<i>kuotama</i>	kuotamii	<i>kerta</i>	kuortoi
Dat	<i>kuotamaa</i>	kuotamazhta	<i>kertaa</i>	kuortozhta
Erg	kuotamuo	kuotamazh	kertuo	kuortozh
All	<i>kuotamaga</i>	kuotamazhka	<i>kertaga</i>	kuortozhka
Ins	<i>kuotamaca</i>	kuotamazhca	<i>kertaca</i>	kuortozhca
Lat	kuotamagh	kuotamegh	kertagh	kuortuojegh
Ins	kuotamal	kuotamel	kertal	kuortuojel
		hen (1)		head (3)

Ingush noun declensions (Nichols 2011)

An abstract stem in Daghestanian?

Two different opinions can be found in the literature: (a) these markers [e.g. -zh JPB] are markers of the ergative case and all oblique cases are formed from the ergative; (b) these markers are markers of the oblique stem (of the singular or plural) and the ergative has no special marker and coincides with the oblique stem of the appropriate number. The first point of view is unsatisfactory: it does not take account of the semantics of the oblique cases (ergative meaning is not a component here), nor of the data from other Daghestanian languages, where the ergative frequently has a special morphological marker like other oblique cases ... (Kibrik 1991: 257)

Oblique stems in Ingush?

- It is straightforward to state the fact that the FORM of genitive singulars and ergative plurals predict the form of the corresponding datives, allatives and instrumentals.
- Expressing these patterns using oblique stems just creates the problem of annotating underspecified stem entries in such a way that they can be 'selected' as the base for the oblique case forms.

Oblique stems in Ingush?

- It is straightforward to state the fact that the FORM of genitive singulars and ergative plurals predict the form of the corresponding datives, allatives and instrumentals.
- Expressing these patterns using oblique stems just creates the problem of annotating underspecified stem entries in such a way that they can be 'selected' as the base for the oblique case forms.

Layered parasitism in Estonian

	Sing	Plu
Nominative	lukk	<i>lukud</i>
Genitive	<i>luku</i>	lukkude
Partitive	lukku	lukkusid
Illa2/Part2	lukku	lukke
Illative	<i>lukusse</i>	lukkudesse
Inessive	<i>lukus</i>	lukkudes
Elative	<i>lukust</i>	lukkudest
Allative	<i>lukule</i>	lukkudele
Adessive	<i>lukul</i>	lukkudel
Ablative	<i>lukult</i>	lukkudelt
Translative	<i>lukuks</i>	lukkudeks
Terminative	<i>lukuni</i>	lukkudeni
Essive	<i>lukuna</i>	lukkudena
Abessive	<i>lukuta</i>	lukkudeta
Comitative	<i>lukuga</i>	lukkudega



Layered structure of case forms

- *lukk* → *lukku* → *lukkude* → *lukkudesse*
 - Nominative singular *lukk* → partitive singular *lukku* → genitive plural *lukkude* → illative plural *lukkudesse*.
 - Stem 1 *lukk* → Stem 2 *lukku* → Stem 3 *lukkude* → *lukkudesse*.
- Stems are identifiable from word forms that they underlie.
- But the stems cannot be assigned properties in isolation that determine their distribution in the paradigm of an item.
- ‘Indexing’ each of these stems in ways that defines their distribution gives rise to a diacritic model of morphology.

Layered structure of case forms

- *lukk* → *lukku* → *lukkude* → *lukkudesse*
 - Nominative singular *lukk* → partitive singular *lukku* → genitive plural *lukkude* → illative plural *lukkudesse*.
 - Stem 1 *lukk* → Stem 2 *lukku* → Stem 3 *lukkude* → *lukkudesse*.
- Stems are identifiable from word forms that they underlie.
- But the stems cannot be assigned properties in isolation that determine their distribution in the paradigm of an item.
- ‘Indexing’ each of these stems in ways that defines their distribution gives rise to a diacritic model of morphology.

Layered structure of case forms

- *lukk* → *lukku* → *lukkude* → *lukkudesse*
 - Nominative singular *lukk* → partitive singular *lukku* → genitive plural *lukkude* → illative plural *lukkudesse*.
 - Stem 1 *lukk* → Stem 2 *lukku* → Stem 3 *lukkude* → *lukkudesse*.
- Stems are identifiable from word forms that they underlie.
- But the stems cannot be assigned properties in isolation that determine their distribution in the paradigm of an item.
- 'Indexing' each of these stems in ways that defines their distribution gives rise to a diacritic model of morphology.

Layered structure of case forms

- *lukk* → *lukku* → *lukkude* → *lukkudesse*
 - Nominative singular *lukk* → partitive singular *lukku* → genitive plural *lukkude* → illative plural *lukkudesse*.
 - Stem 1 *lukk* → Stem 2 *lukku* → Stem 3 *lukkude* → *lukkudesse*.
- Stems are identifiable from word forms that they underlie.
- But the stems cannot be assigned properties in isolation that determine their distribution in the paradigm of an item.
- ‘Indexing’ each of these stems in ways that defines their distribution gives rise to a diacritic model of morphology.

Layered structure of case forms

- *lukk* → *lukku* → *lukkude* → *lukkudesse*
 - Nominative singular *lukk* → partitive singular *lukku* → genitive plural *lukkude* → illative plural *lukkudesse*.
 - Stem 1 *lukk* → Stem 2 *lukku* → Stem 3 *lukkude* → *lukkudesse*.
- Stems are identifiable from word forms that they underlie.
- But the stems cannot be assigned properties in isolation that determine their distribution in the paradigm of an item.
- 'Indexing' each of these stems in ways that defines their distribution gives rise to a diacritic model of morphology.

Layered structure of case forms

- *lukk* → *lukku* → *lukkude* → *lukkudesse*
 - Nominative singular *lukk* → partitive singular *lukku* → genitive plural *lukkude* → illative plural *lukkudesse*.
 - Stem 1 *lukk* → Stem 2 *lukku* → Stem 3 *lukkude* → *lukkudesse*.
- Stems are identifiable from word forms that they underlie.
- But the stems cannot be assigned properties in isolation that determine their distribution in the paradigm of an item.
- ‘Indexing’ each of these stems in ways that defines their distribution gives rise to a diacritic model of morphology.

Implicational structure of grammatical case forms

- The paradigm of LUKK is predictable from partitive singular *lukku*.
- The grammatical forms are directly predictable via
 - Identity: 'short' illative singular *lukku*,
 - Truncation: nominative singular *lukk*,
 - Shortening: genitive singular *luku*,
 - Suffixation: genitive plurals *lukkude* and partitive plural *lukkusid*.

Implicational structure of grammatical case forms

- The paradigm of LUKK is predictable from partitive singular *lukku*.
- The grammatical forms are directly predictable via
 - Identity: 'short' illative singular *lukku*,
 - Truncation: nominative singular *lukk*,
 - Shortening: genitive singular *luku*,
 - Suffixation: genitive plurals *lukkude* and partitive plural *lukkusid*.

Implicational structure of grammatical case forms

- The paradigm of LUKK is predictable from partitive singular *lukku*.
- The grammatical forms are directly predictable via
 - Identity: 'short' illative singular *lukku*,
 - Truncation: nominative singular *lukk*,
 - Shortening: genitive singular *luku*,
 - Suffixation: genitive plurals *lukkude* and partitive plural *lukkusid*.

Implicational structure of grammatical case forms

- The paradigm of LUKK is predictable from partitive singular *lukku*.
- The grammatical forms are directly predictable via
 - Identity: 'short' illative singular *lukku*,
 - Truncation: nominative singular *lukk*,
 - Shortening: genitive singular *luku*,
 - Suffixation: genitive plurals *lukkude* and partitive plural *lukkusid*.



Implicational structure of grammatical case forms

- The paradigm of LUKK is predictable from partitive singular *lukku*.
- The grammatical forms are directly predictable via
 - Identity: 'short' illative singular *lukku*,
 - Truncation: nominative singular *lukk*,
 - Shortening: genitive singular *luku*,
 - Suffixation: genitive plurals *lukkude* and partitive plural *lukkusid*.

Implicational structure of grammatical case forms

- The paradigm of LUKK is predictable from partitive singular *lukku*.
- The grammatical forms are directly predictable via
 - Identity: 'short' illative singular *lukku*,
 - Truncation: nominative singular *lukk*,
 - Shortening: genitive singular *luku*,
 - Suffixation: genitive plurals *lukkude* and partitive plural *lukkusid*.

Implicational structure of semantic case forms

- The endings of the Illative through Comitative are invariant.
- Hence the significant variation lies in stem selection.
- Semantic forms are based on the corresponding genitive:
 - Singular forms are based on genitive singular *luke*,
 - (The nominative singular is also based on the genitive singular),
 - Plural forms are based on genitive plural *lukkude*.
- But since the case endings are invariant, implications are SYMMETRICAL and semantic case forms imply their genitive bases.

Implicational structure of semantic case forms

- The endings of the Illative through Comitative are invariant.
- Hence the significant variation lies in stem selection.
- Semantic forms are based on the corresponding genitive:
 - Singular forms are based on genitive singular *luke*,
 - (The nominative singular is also based on the genitive singular),
 - Plural forms are based on genitive plural *lukkude*.
- But since the case endings are invariant, implications are SYMMETRICAL and semantic case forms imply their genitive bases.

Implicational structure of semantic case forms

- The endings of the Illative through Comitative are invariant.
- Hence the significant variation lies in stem selection.
- Semantic forms are based on the corresponding genitive:
 - Singular forms are based on genitive singular *luke*,
 - (The nominative singular is also based on the genitive singular),
 - Plural forms are based on genitive plural *lukkude*.
- But since the case endings are invariant, implications are SYMMETRICAL and semantic case forms imply their genitive bases.

Implicational structure of semantic case forms

- The endings of the Illative through Comitative are invariant.
- Hence the significant variation lies in stem selection.
- Semantic forms are based on the corresponding genitive:
 - Singular forms are based on genitive singular *luke*,
 - (The nominative singular is also based on the genitive singular),
 - Plural forms are based on genitive plural *lukkude*.
- But since the case endings are invariant, implications are SYMMETRICAL and semantic case forms imply their genitive bases.

Implicational structure of semantic case forms

- The endings of the Illative through Comitative are invariant.
- Hence the significant variation lies in stem selection.
- Semantic forms are based on the corresponding genitive:
 - Singular forms are based on genitive singular *luke*,
 - (The nominative singular is also based on the genitive singular),
 - Plural forms are based on genitive plural *lukkude*.
- But since the case endings are invariant, implications are SYMMETRICAL and semantic case forms imply their genitive bases.

Implicational structure of semantic case forms

- The endings of the Illative through Comitative are invariant.
- Hence the significant variation lies in stem selection.
- Semantic forms are based on the corresponding genitive:
 - Singular forms are based on genitive singular *luke*,
 - (The nominative singular is also based on the genitive singular),
 - Plural forms are based on genitive plural *lukkude*.
- But since the case endings are invariant, implications are SYMMETRICAL and semantic case forms imply their genitive bases.

Domains of implication

- An implicational analysis captures the fact that the stem variation exhibited by Ingush and Estonian does not signal semantic properties but instead sanctions deductions about other forms.
- In Estonian, grammatical case forms, singular semantic forms and plural semantic forms each comprise interpredictable cohort sets. The same factorization applies to more familiar systems.
- Implicational analyses offer revealing factorizations of other systems by identifying domains over which implications apply.

Domains of implication

- An implicational analysis captures the fact that the stem variation exhibited by Ingush and Estonian does not signal semantic properties but instead sanctions deductions about other forms.
- In Estonian, grammatical case forms, singular semantic forms and plural semantic forms each comprise interpredictable cohort sets. The same factorization applies to more familiar systems.
- Implicational analyses offer revealing factorizations of other systems by identifying domains over which implications apply.

Domains of implication

- An implicational analysis captures the fact that the stem variation exhibited by Ingush and Estonian does not signal semantic properties but instead sanctions deductions about other forms.
- In Estonian, grammatical case forms, singular semantic forms and plural semantic forms each comprise interpredictable cohort sets. The same factorization applies to more familiar systems.
- Implicational analyses offer revealing factorizations of other systems by identifying domains over which implications apply.

How many declension classes in German?

In reality, the choice of plural formation depends largely on gender and/or inflection class as manifested also in the expression of the four German cases in the singular. Thus if a masculine has the suffix -en in the Gen.Sg., it must also have it for the plural, e.g. der Fürst 'prince, sovereign', Gen.Sg. des Fürst-en implies the plural Fürst-en. (Laaha ea 2006: 279)

Singular patterns

	Masc		Neut	Fem
	S1	S2	S1	S3
Nom	Pegel	Prinz	Segel	Regel
Acc	Pegel	Prinzen	Segel	Regel
Dat	Pegel	Prinzen	Segel	Regel
Gen	Pegels 'level'	Prinzen 'bear'	Segels 'sail'	Regel 'rule'

Singular declensional patterns in German (cf. Duden (2005: 197))

Plural patterns

Ending	P1	P2		P3	P4		P5
Stem	-s -uml	-(e)n -uml	-e -uml	-e +uml	-er +uml	∅ -uml	∅ +uml
N/A/G	Uhus	Prinzen	Hunde	Bünde	Münder	Balken	Gärten
Dat (Masc)	Uhus 'owl'	Prinzen 'bear'	Hunden 'dog'	Bünden 'waistcoat'	Mündern 'mouth'	Balken 'beam'	Gärten 'garden'
N/A/G	Autos	Ohren	Jahre	Flöße	Länder	Muster	Klöster
Dat (Neut)	Autos 'cars'	Ohren 'eyes'	Jahren 'years'	Flößen 'rafts'	Ländern 'countries'	Mustern 'patterns'	Klöstern 'cloisters'
N/A/G	Bars	Regeln	—	Hände	—	—	Töchter
Dat (Fem)	Bars 'bars'	Regeln 'rules'	—	Händen 'hands'	—	—	Töchtern 'daughters'

Plural declensional patterns in German (cf. Duden (2005: 226))

Declensional patterns

	uml	Masc	S1 Neut	S2 Masc	S3 Fem
P1	–	UHU	AUTO	–	KAMERA
P2	–	STAAT	OHR	PRINZ	REGEL
P3	–	HUND	JAHR	–	–
P3	+	BUND	(FLOSS)	–	HAND
P4	+	MUND	LAND	–	–
P5	–	BALKEN	MUSTER	–	–
P5	+	GARTEN	(KLOSTER)	–	TOCHTER

Combinations of singular and plural patterns in German

Structure of German noun declensions

- Singular forms comprise a cohort set, in which the genitive is the most informative form. Plural forms also comprise a cohort set.
- There is a weaker implicational relation between singular and plural forms, conditioned by gender and phonotactics:
 - $S_2 \rightarrow P_2$ ('weak' masculine singulars are weak in the plural).
 - $S_3 \rightarrow \neg P_4$ (no feminine plurals in *-er*).
 - S_1 ('strong' masculine/neuter) is compatible with all plurals.

Structure of German noun declensions

- Singular forms comprise a cohort set, in which the genitive is the most informative form. Plural forms also comprise a cohort set.
- There is a weaker implicational relation between singular and plural forms, conditioned by gender and phonotactics:
 - $S_2 \rightarrow P_2$ ('weak' masculine singulars are weak in the plural).
 - $S_3 \rightarrow \neg P_4$ (no feminine plurals in *-er*).
 - S_1 ('strong' masculine/neuter) is compatible with all plurals.

Structure of German noun declensions

- Singular forms comprise a cohort set, in which the genitive is the most informative form. Plural forms also comprise a cohort set.
- There is a weaker implicational relation between singular and plural forms, conditioned by gender and phonotactics:
 - $S_2 \rightarrow P_2$ ('weak' masculine singulars are weak in the plural).
 - $S_3 \rightarrow \neg P_4$ (no feminine plurals in *-er*).
 - S_1 ('strong' masculine/neuter) is compatible with all plurals.

Structure of German noun declensions

- Singular forms comprise a cohort set, in which the genitive is the most informative form. Plural forms also comprise a cohort set.
- There is a weaker implicational relation between singular and plural forms, conditioned by gender and phonotactics:
 - $S_2 \rightarrow P_2$ ('weak' masculine singulars are weak in the plural).
 - $S_3 \rightarrow \neg P_4$ (no feminine plurals in *-er*).
 - S_1 ('strong' masculine/neuter) is compatible with all plurals.

Structure of German noun declensions

- Singular forms comprise a cohort set, in which the genitive is the most informative form. Plural forms also comprise a cohort set.
- There is a weaker implicational relation between singular and plural forms, conditioned by gender and phonotactics:
 - $S_2 \rightarrow P_2$ ('weak' masculine singulars are weak in the plural).
 - $S_3 \rightarrow \neg P_4$ (no feminine plurals in *-er*).
 - S_1 ('strong' masculine/neuter) is compatible with all plurals.

Measuring predictive value

- Let paradigm cells be random variables that take realization 'outcomes' as their values (e.g., for [Gen Sg] the value 'Xs').
- The uncertainty associated with the realization of a cell C can be defined in terms of the ENTROPY (Shannon 1948) of the cell, $H(C)$:

$$H(C) = - \sum_{x \in R_C} p(x) \log_2 p(x)$$

- R_C here represents the set of realization outcomes for C , x an outcome in R , and $p(x)$ the probability that C is realized by x .

Measuring predictive value

- Let paradigm cells be random variables that take realization 'outcomes' as their values (e.g., for [Gen Sg] the value 'Xs').
- The uncertainty associated with the realization of a cell C can be defined in terms of the ENTROPY (Shannon 1948) of the cell, $H(C)$:

$$H(C) = - \sum_{x \in R_C} p(x) \log_2 p(x)$$

- R_C here represents the set of realization outcomes for C , x an outcome in R , and $p(x)$ the probability that C is realized by x .

Measuring predictive value

- Let paradigm cells be random variables that take realization 'outcomes' as their values (e.g., for [Gen Sg] the value 'Xs').
- The uncertainty associated with the realization of a cell C can be defined in terms of the ENTROPY (Shannon 1948) of the cell, $H(C)$:

$$H(C) = - \sum_{x \in R_C} p(x) \log_2 p(x)$$

- R_C here represents the set of realization outcomes for C , x an outcome in R , and $p(x)$ the probability that C is realized by x .

Morphological information

- The diagnostic value of an individual cell correlates with the degree to which it reduces uncertainty about other cells.
- Uncertainty reduction can be expressed in terms of **CONDITIONAL ENTROPY**, $H(C_2|C_1)$, which measures the amount of uncertainty that remains about C_2 given knowledge of C_1 .
- **MORPHOLOGICAL INFORMATION** measures uncertainty reduction:

$$M(C_2|C_1) = 1 - \frac{H(C_2|C_1)}{H(C_2)}$$

Morphological information

- The diagnostic value of an individual cell correlates with the degree to which it reduces uncertainty about other cells.
- Uncertainty reduction can be expressed in terms of **CONDITIONAL ENTROPY**, $H(C_2|C_1)$, which measures the amount of uncertainty that remains about C_2 given knowledge of C_1 .
- **MORPHOLOGICAL INFORMATION** measures uncertainty reduction:

$$M(C_2|C_1) = 1 - \frac{H(C_2|C_1)}{H(C_2)}$$

Morphological information

- The diagnostic value of an individual cell correlates with the degree to which it reduces uncertainty about other cells.
- Uncertainty reduction can be expressed in terms of **CONDITIONAL ENTROPY**, $H(C_2|C_1)$, which measures the amount of uncertainty that remains about C_2 given knowledge of C_1 .
- **MORPHOLOGICAL INFORMATION** measures uncertainty reduction:

$$\mathbb{M}(C_2|C_1) = 1 - \frac{H(C_2|C_1)}{H(C_2)}$$

Paradigm uncertainty

- The cumulative uncertainty associated with a paradigm \mathcal{P} depends directly on the uncertainty of its cells C_1, C_2, \dots, C_n .
- On a traditional model, cells are generally assumed to be interdependent, so that the entropy of a paradigm, $H(\mathcal{P})$, will correspond to the JOINT ENTROPY of its cells, $H(C_1, C_2, \dots, C_n)$.
- The information that C expresses about a paradigm \mathcal{P} is then

$$\mathbb{M}(\mathcal{P}|C) = 1 - \frac{H(\mathcal{P}|C)}{H(\mathcal{P})}$$

Paradigm uncertainty

- The cumulative uncertainty associated with a paradigm \mathcal{P} depends directly on the uncertainty of its cells C_1, C_2, \dots, C_n .
- On a traditional model, cells are generally assumed to be interdependent, so that the entropy of a paradigm, $H(\mathcal{P})$, will correspond to the JOINT ENTROPY of its cells, $H(C_1, C_2, \dots, C_n)$.
- The information that C expresses about a paradigm \mathcal{P} is then

$$\mathbb{M}(\mathcal{P}|C) = 1 - \frac{H(\mathcal{P}|C)}{H(\mathcal{P})}$$

Paradigm uncertainty

- The cumulative uncertainty associated with a paradigm \mathcal{P} depends directly on the uncertainty of its cells C_1, C_2, \dots, C_n .
- On a traditional model, cells are generally assumed to be interdependent, so that the entropy of a paradigm, $H(\mathcal{P})$, will correspond to the JOINT ENTROPY of its cells, $H(C_1, C_2, \dots, C_n)$.
- The information that C expresses about a paradigm \mathcal{P} is then

$$\mathbb{M}(\mathcal{P}|C) = 1 - \frac{H(\mathcal{P}|C)}{H(\mathcal{P})}$$

Principal parts

- The implicational relations between cells invoked by traditional grammars can be modelled by morphological information.
- A cell or set of cells is **DIAGNOSTIC** if its morphological information value approaches 1.
- Cells are highly nondiagnostic when their value approaches 0.

Principal parts

- The implicational relations between cells invoked by traditional grammars can be modelled by morphological information.
- A cell or set of cells is **DIAGNOSTIC** if its morphological information value approaches 1.
- Cells are highly nondiagnostic when their value approaches 0.

Principal parts

- The implicational relations between cells invoked by traditional grammars can be modelled by morphological information.
- A cell or set of cells is **DIAGNOSTIC** if its morphological information value approaches 1.
- Cells are highly nondiagnostic when their value approaches 0.

Principal parts

- This provides one solution to the traditional problem of justifying principal parts (cf. also Finkel & Stump 2008, 2009)

One objection to the Priscianic model ... was that the choice of leading form was inherently arbitrary: the theory creates a problem which it is then unable, or only partly able, to resolve (Matthews 1972: 74).

Principal parts

- This provides one solution to the traditional problem of justifying principal parts (cf. also Finkel & Stump 2008, 2009)

One objection to the Priscianic model ... was that the choice of leading form was inherently arbitrary: the theory creates a problem which it is then unable, or only partly able, to resolve (Matthews 1972: 74).

Summary

- Word and paradigm approaches are item and pattern models.
 - The patterns provide a base for analogical deduction.
 - Words are the most informative items in a morphological system.
 - Inflectional paradigms exhibit reliable implicational patterns.
- Implicational analyses of morphological systems model a type of information that speakers need for language acquisition and use.

Summary

- Word and paradigm approaches are item and pattern models.
 - The patterns provide a base for analogical deduction.
 - Words are the most informative items in a morphological system.
 - Inflectional paradigms exhibit reliable implicational patterns.
- Implicational analyses of morphological systems model a type of information that speakers need for language acquisition and use.

Summary

- Word and paradigm approaches are item and pattern models.
 - The patterns provide a base for analogical deduction.
 - Words are the most informative items in a morphological system.
 - Inflectional paradigms exhibit reliable implicational patterns.
- Implicational analyses of morphological systems model a type of information that speakers need for language acquisition and use.

Summary

- Word and paradigm approaches are item and pattern models.
 - The patterns provide a base for analogical deduction.
 - Words are the most informative items in a morphological system.
 - Inflectional paradigms exhibit reliable implicational patterns.
- Implicational analyses of morphological systems model a type of information that speakers need for language acquisition and use.

Summary

- Word and paradigm approaches are item and pattern models.
 - The patterns provide a base for analogical deduction.
 - Words are the most informative items in a morphological system.
 - Inflectional paradigms exhibit reliable implicational patterns.
- Implicational analyses of morphological systems model a type of information that speakers need for language acquisition and use.