UNIVERSITY OF OSLO

The source of variation in language

Sverre Stausland
sverre.stausland@gmail.com

May 15, 2024

Source of variation

- By the 'source' of variation we can mean two distinct things:
  - The diachronic source
  - The synchronic (or grammatical or mental) source

- I will try to say something about both.

Definition of variation

- *Variation* here will mean that
  - a speaker uses different forms to express the 'same' meaning
  - all forms belong to what the speaker considers to be the 'same' language

- In this talk I will limit the discussion to morphology and phonology.

Variation in morphology from change

- In morphology, I will assume that no original state has two alternative forms with the same meaning.

- When such a state exists, one form is original and one is new.

- The new form is a result of *change*.

- Variation in morphology is thus the result of language change.

Morphological change 1 – grammaticalization

- A function word can become an affix through coalescence with an adjacent word.

- In Scandinavian, the third person reflexive pronoun *sik* became a verbal suffix *-sk* expressing mediopassive.

- In Old and Modern Norwegian, many verbs can either use the reflexive pronoun (original) or a mediopassive suffix (new), with no apparent difference in meaning.

Morphological change 1 – grammaticalization

- *breida seg* ~ *breidast* 'disperse'

- *brigda seg* ~ *brigdast* 'change'

- *eva seg* ~ *evast* 'doubt'

Morphological change 2 – analogy

- A declension pattern can extend beyond its original domain.

- In East Norwegian, *-i* is a salient marker of the perfect in strong verbs.

- It has spread to weak verbs.

- In weak verbs, the original perfect endings *-a* and *-t* now exist alongside the new ending *-i*.

Morphological change 2 – analogy

- *bada* ~ *badi* 'bathed'

- *svingt* ~ *svingi* 'turned'

Morphological change 2 – analogy

- In Old English, an original unstressed *ō developed to *u* in medial syllables and *a* in final syllables (Stausland Johnsen 2015).

- This would give alternations such as *wundude – wundad* 'wounded'.

- In analogy with other verbal paradigms where no such vowel alternations existed, either *u* or *a* could be generalized:

  wundude   ~   *wundud*

  wundade   ~   *wundad*

Morphological change 3 – loss of contrast

- A morphological contrast expressed with different endings can be lost.

- After the loss of the contrast, the two endings can be used interchangeably.

- East Norwegian dialects distinguish between dative and 'non-dative' cases in nouns.

- *-om* (dative plural) vs. *-ne* (non-dative plural)

Morphological change 3 – loss of contrast

- This contrast is now lost for many speakers.

- The original case forms *-om* and *-ne* can now be used interchangeably to express the plural.

- Real example (NRK Dagsnytt 18, 4/4 2023):

  Hytt-om kan bli brukt mer – Å bruke dei hytt-ne vi allerede har bygd

Morphological change 4 – language contact

- When language varieties meet, morphemes can be borrowed and used interchangeably with the original morphemes.

- For many Norwegian speakers, the original past tense suffix *-a* can be used interchangeably with the borrowed (Norwegian-)Danish suffix *-et*.

- *kasta* ~ *kastet* 'thrown'

- *skuffa* ~ *skuffet* 'disappointed'

Variation in phonology

- Variation in phonology is either:
  - A phonological process is optional, resulting in two possible outcomes.
  - Two or more ways of realizing a phoneme in the same context.

- Unlike morphology, phonological variation is not necessarily the result of historical changes.

Phonological change 1 – sound change

- I will assume that all synchronic phonological processes are fossilized sound changes.

- At some point in the history of East Norwegian, the sequence /rt/ assimilated to a retroflex segment /ʈ/ (Stausland Johnsen 2012).

- This sound change is fossilized as a synchronic phonological rule that takes /r#t/ across morpheme boundaries to /ʈ/:

- /føːr-/ 'lead' + /-te/ 'PRET' > /føːʈe/

  /stuːr-/ 'big' + /-t/ 'NEUT' > /stuːʈ/

Phonological change 1 – sound change

- Some phonological processes of this kind are optional.

- An /n/ assimilates to the place features of a following consonant in Norwegian.

- /sæjn/ 'slow' (adj.) → [seŋ-kə] 'slow' (verb)

  /grøːn/ 'green' → [grøŋ-kɑs] 'become green'

Phonological change 1 – sound change

- Across lexical boundaries, however, this process is optional.

- /san/ 'sand' + /kaːke/ 'cake' > [sɑnkɑːkə] ~ [sɑŋkɑːkə] 'cookie'

  /bran/ 'fire' + /biːl/ 'car' > [brɑnbiːl] ~ [brɑmbiːl] 'fire truck'

- This gives variation in the realization of /n/ in the same words.

Phonological change 2 – loss of contrast

- When the contrast between two phonemes is lost, speakers could end up using them interchangeably in the same contexts.

- The contrast between /ç/ and /ʃ/ is being lost in Norwegian.

- It has been reported (anecdotally) that some speakers now use [ç] and [ʃ] interchangeably as the realization of the same phoneme.

Phonological change 3 – language contact

- When language varieties meet, their phonological systems and inventories can influence each other.

- In East Norwegian, /l/ and /ɽ/ are two separate phonemes.

- In Norwegian-Danish, on the other hand, /ɽ/ is not part of the phonemic inventory.

- East Norwegian words with /ɽ/ have /l/ in Norwegian-Danish.

Phonological change 3 – language contact

- Contact between these varieties has caused many speakers of East Norwegian to realize the phoneme /ɽ/ variably as [ɽ] or [l].

- [bɽɔː]   ~   [blɔː] 'blue'

  [suːɽ]   ~   [suːl] 'sun'

Phonological variation without change

- At least under one condition, phonological variation exists without there having been a change to cause it.

- The realization of a phoneme often varies along a dimension there is no contrast for.

- In word initial position in Norwegian, aspirated stops contrast with non-aspirated stops.

- There is no voice contrast.

Phonological variation without change

- Initial non-aspirated stops vary between a voiced and a voiceless realization.

- /buː/ = [buː] ~ [puː] 'live'

Phonological variation without change

- Norwegian vowels contrast in length: short and long vowels.

- But there is no contrast between long vowels and 'half-long' or 'super-long' vowels.

- Long vowels in Norwegian vary considerably in how long they are.

- [mɑˑt] ~ [mɑːt] ~ [mɑːːt]

Variation in the grammar

- These are some of the many ways from which variation can arise in a language.

- But how is such variation represented in the mental grammar of a speaker?

- And how does a speaker 'choose' which variant to produce in a given context?

Probabilistic grammars

- One traditional way of encoding variation is with different kinds of probabilistic grammars.

- Probabilities are hard coded into the grammar.

- Say it is observed that a speaker produces variant A 25 % of the time and variant B 75 % of the time.

- These probabilities are then hard coded into the grammar of this speaker (the implementation varies by model).

- This will ensure that when a form is produced, there is a 25 % chance of producing A, and a 75 % chance of producing B.

Probabilistic grammars

- The typical problem with such approaches is that they do not explain anything.

- Instead, they are formalizations of observations.

- I.e. we observe X in the output, so we put X in the grammar to ensure that the speaker produces X.

Probabilistic grammars

- More relevant here is that a purely probabilistic approach assumes that the only thing distinguishing the variants is their frequencies.

- In other words, the choice of one variant over another has nothing to do with the context (linguistic or extra-linguistic).

- This is probably incorrect.

Factors influencing choice of variant

- We know of many factors which strongly influence which variant is 'chosen'.

- They include:
    - Speech rate
    - Interlocutor or audience
    - Social setting
    - Theme (what you are talking about)

Probabilistic grammars

- Models of probabilistic grammars can implement mechanisms in which these factors change the probabilities of different variants.

- E.g. when you shift to an interlocutor of a specific type, the probabilities of variant A vs. variant B changes from 25 % vs. 75 % to 65 % vs. 35 %.

- But this is also just hard coding the observations, and there is no explanation for why these factors change the probabilities the way they do.

Exemplar model

- One model that has been quite successful in explaining variation in phonetics is the "exemplar model".

- Here I will lay out the core properties of the model and suggest how it can account for variation in phonology and morphology as well.

Exemplar model

- It was long an assumption that phonetic and extra-linguistic details of a perceived word were 'stripped away' as the word was passed through the perceptual and representational systems.

- E.g. when a listener perceived the word *cat*, all that remained in the end was the phonemic string /kæt/.

- A range of experiments have shown that this is not true.

Exemplar model

- When perceiving a pronunciation of the word *cat*, listeners actually 'remember':
  - □ Phonetic details (length, pitch, aspiration, vowel quality, etc.)
  - □ Extra-linguistic details (voice, loudness, gender, situation, etc.)

- Perceived tokens of a word are stored as individual *episodic memories*, which we call *exemplars*.

Evidence for exemplars

- In several experiments, people have been recorded reading out lists of words.

- Then they listen to recordings of other people pronouncing these words.

- Then they are asked to read out the words again.

- In the second round of reading, their pronunciations have shifted significantly towards the pronunciations they had heard from others.

Evidence for exemplars

- Somehow the phonetic details of what they have heard are influencing what they themselves produce.

- This is not possible unless these phonetic details have been stored somehow.

- Since features from these episodes then show up in their own productions, it means that speakers draw from these exemplars when producing speech.

- Stored episodic memories (exemplars) form the basis for linguistic production.

Exemplar model

- The mental representation of a word consists of a 'cloud' of exemplars.

- In production, speakers draw from this cloud.

- Most people assume some noise/randomness and averaging over exemplars in the selection procedure.

Exemplar model

- Exemplars are 'tagged' with different features.

- All exemplars produced by men are tagged 'male', exemplars produced in a formal setting are tagged 'formal', etc.

- Clouds are connected by tags in a multidimensional network.

Exemplar model

- All exemplars in the cloud are not equal.

- They differ in terms of *activation*, just like other memories do.

- Recent exemplars are more activated, while old exemplars weaken and fade away (as memories do).

Exemplar model

- Exemplars can be activated through their tags and connections.

- Exemplars which were perceived in formal settings, are tagged 'formal'.

- When the speaker is in a formal setting, exemplars tagged 'formal' are activated.

- Activated exemplars will in turn activate other exemplars they are connected with.

Exemplar model

- When the speaker produces a word, the more activated exemplars contribute more to the selection process.

- Speakers will therefore mimic recently heard words more than words heard a long time ago.

- And speakers will mimic words produced in similar contexts as the words they are currently uttering.

Evidence for the exemplar cloud

- The cloud for a word consists of episodic memories of that word being produced.

- The more frequent a word is, the more exemplars in the cloud.

- When a word is perceived and stored, its contribution to the cloud is inversely proportional to the size of the cloud.

- The more exemplars a word has, the less a new exemplar will contribute to its cloud overall.

Evidence for the exemplar cloud

- As a result, we expect to find that speakers are less influenced by recently perceived tokens for high frequency words than for low frequency words.

- Speakers should mimic newly perceived words more when these words have low frequency.

- This is exactly what experiments show.

Exemplar model and variation – phonology

- Speakers have exemplars tagged with /ç/ and /ʃ/.

- Some speakers have lost this contrast, and so they now all have the same tag, e.g. /ʃ/.

- But the exemplars still contain [ç] and [ʃ], since this is what was being perceived.

- When the speaker wants to produce /ʃ/, all exemplars with this tag are activated, thus activating both [ç] and [ʃ].

- The speaker will then produce both [ç] and [ʃ] as variants of /ʃ/.

Exemplar model and variation – phonology

- Speakers have a cloud for the word *sol*.

- Tokens produced by members of the upper class or in formal settings will be tagged accordingly, others as lower class or informal.

- When the speaker finds himself in a formal setting, the tag 'formal' is activated, thus activating more exemplars of the kind [suːl], making this variant more likely to be uttered.

- Conversely, in an informal setting, 'informal' is activated, making [suːɽ] more likely.

Exemplar model and variation – morphology

- A speaker wants to produce the perfect form of a weak verb.

- The perfect forms of that word will thus be activated.

- But these exemplars are connected with other verbs also tagged with 'PERF'.

- The perfect forms in *-i* of strong verbs will also be activated.

- As a result, the *-i* from a strong verb will occasionally be selected instead of the 'correct' weak verb ending *-a*.

Exemplar model and variation – morphology

- In dialects with a dative vs. non-dative contrast, the plural ending *-om* is tagged 'DAT.PL.', and the plural ending *-ne* with 'NON-DAT.PL'.

- For some speakers, the case contrast is lost, so that all plural forms are simply tagged 'PL'.

Exemplar model and variation – morphology

- The clouds for these plural forms will contain both forms in *-om* and *-ne*.

- When such a speaker wants to produce a plural form, *-om* and *-ne* are both activated.

- The speaker will then produce both *-om* and *-ne* as variants of the plural.

Exemplar model and variation

- These were some illustrations of how the exemplar model can be useful in explaining how and why variation exists in a synchronic grammar.

- Whether this can be extended to all cases of synchronic variation, is less clear.

- While the exemplar model has been a popular model to explain phonetic production, no one has (to my knowledge) tried to incorporate an exemplar framework into a theory of grammar.

- It is even less clear how the exemplar model can be applied to syntax.

References I

📄 Stausland Johnsen, Sverre (Dec. 2012). A diachronic account of phonological unnaturalness. *Phonology* 29.3: 505–531. DOI: 10.1017/S0952675712000243.

📄 — (2015). Vowel reduction in verbs in King Alfred's Pastoral Care. *Historical linguistics 2013. Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5–9 August 2013*. Ed. by Dag T. T. Haug. Current issues in linguistic theory. Amsterdam studies in the theory and history of linguistic science – series IV 334. Amsterdam: John Benjamins, 19–40. DOI: 10.1075/cilt.334.02joh.