MUNI
ARTS

Morphological productivity

Guy Tabachnick
gtabach.github.io

Faculty of Arts, Masaryk University

OVA 2026

Introduction

One of the major aspects of knowing a language is knowing how to *build words*.

- How do English speakers know that the past tense of *go* is *went*?
- How did English speakers agree that the plural of *google* should be *googled*?

Productivity involves the latter question: how do speakers build *new words* when simple lookup of stored forms is not an option?

- By studying productivity, we can learn about speakers' knowledge of *abstract patterns* about their language
- This gives us crucial insight into the nature of the grammar

Course overview

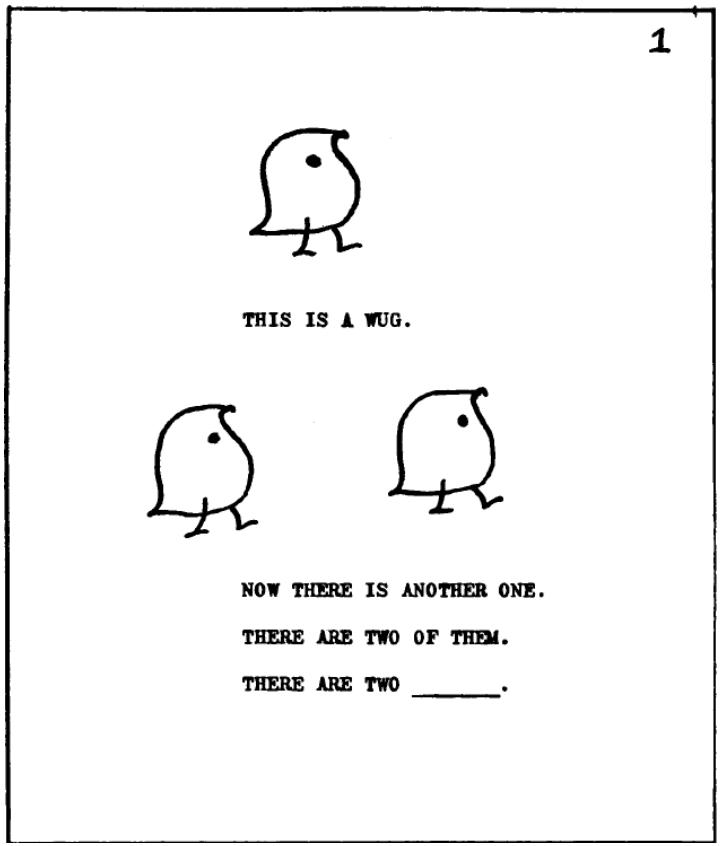
- What is productivity?
- How do we test productivity?
- How do we model productivity?
- What patterns do speakers use productively?

Outline

The wug test

Berko (1958) famously invented the *wug test* (often known today as the *nonce word test*) to test productivity:

"In this study we set out to discover what is learned by children exposed to English morphology. To test for knowledge of morphological rules, we use nonsense materials. [...] If a child knows that the plural of witch *is* witches, *he may simply have memorized the plural form. If, however, he tells us that the plural of* *gutch *is* *gutches, *we have evidence that he actually knows, albeit unconsciously, one of those rules which the descriptive linguist, too, would set forth in his grammar."* (Berko, 1958, p. 150)



English plurals

The vast majority of English nouns take one of three plural forms depending on the final consonant of the noun:

- [əz] after sibilants [s z ʃ ʒ tʃ dʒ] (*glasses*, *watches*)
- [s] after voiceless consonants [p t k f θ] (*raps*, *bits*)
- [z] after all other consonants [b d g v ð m n ŋ ɹ l w j] and vowels (*lids*, *bows*)

On rare occasions, the plural is formed otherwise (different *allomorphs*): *men*, *women*, *oxen*, *sheep*

Adult performance

Adults were usually unanimous:

	singular	*plural*
‘wug’	wəg	wəgz
‘heaf’	hijf	hijfs / hijvz
‘lun’	lən	lənz
‘tor’	tɔɹ	tɔɹz
‘cra’	krɑ	krɑz
‘tass’	tæs	tæsəz
‘gutch’	gətʃ	gətʃəz
‘kazh’	kæʒ	kæʒəz
‘niz’	nɪz	nɪzəz
‘glass’	glæs	glæsəz

About half said [hijvz], following a fairly common pattern in English: [najf] → [najvz] but [rijf] → [rijvz]

Child performance

Children ages 4–7 varied:

	"correct"	*% "correct"*	
singular	*plural*	*4–5*	*5.5–7*
wəg	wəgz	76	97
hijf	hijfs/vz	79	80
lən	lənz	68	92
tɔɹ	tɔɹz	73	90
krɑ	krɑz	58	86
tæs	tæsəz	28	39
gətʃ	gətʃəz	28	38
kæʒ	kæʒəz	25	36
nɪz	nɪzəz	14	33
glæs	glæsəz	75	99

Child performance

Children ages 4–7 varied:

	"correct"	*% "correct"*	
singular	*plural*	*4–5*	*5.5–7*
wəg	wəgz	76	97
hijf	hijfs/vz	79	80
lən	lənz	68	92
tɔɹ	tɔɹz	73	90
krɑ	krɑz	58	86
tæs	tæsəz	28	39
gətʃ	gətʃəz	28	38
kæʒ	kæʒəz	25	36
nɪz	nɪzəz	14	33
glæs	glæsəz	75	99

- By age 6, children have pretty much figured out [z] and [s]

Child performance

Children ages 4–7 varied:

	"correct"	*% "correct"*	
singular	*plural*	*4–5*	*5.5–7*
wəg	wəgz	76	97
hijf	hijfs/vz	79	80
lən	lənz	68	92
tɔɹ	tɔɹz	73	90
krɑ	krɑz	58	86
tæs	tæsəz	28	39
gətʃ	gətʃəz	28	38
kæʒ	kæʒəz	25	36
nɪz	nɪzəz	14	33
glæs	glæsəz	75	99

- By age 6, children have pretty much figured out [z] and [s]
- They have also learned the plural of common words ending in sibilants

Child performance

Children ages 4–7 varied:

	"correct"	*% "correct"*	
singular	*plural*	*4–5*	*5.5–7*
wəg	wəgz	76	97
hijf	hijfs/vz	79	80
lən	lənz	68	92
tɔɹ	tɔɹz	73	90
krɑ	krɑz	58	86
tæs	tæsəz	28	39
gətʃ	gətʃəz	28	38
kæʒ	kæʒəz	25	36
nɪz	nɪzəz	14	33
glæs	glæsəz	75	99

- By age 6, children have pretty much figured out [z] and [s]
- They have also learned the plural of common words ending in sibilants
- But they continue to struggle with these words, confidently preferring plurals like [tæs] and sometimes [gətʃs]

Summary

- There's a difference between knowing the plural of a word and knowing how to *productively* extend plural-making principles to new words
- Adults apply all three of the "regular" processes, basically without exception
- Some of the patterns take longer for children to learn than others
- By manipulating properties of the target words, we can test speakers' understanding of different patterns

Outline

Acquisition studies

The wug test can tell us a lot about how children learn language, but it has some issues:

- It's an artificial task, so people may do different things than when they're speaking naturally
- Every child learns at slightly different rates, so getting many children at a single point in time can be unclear

An alternative is to track individuals over longer periods of time

- Children come into the lab at regular intervals
- The lab has a bunch of toys and their interactions with caretakers are recorded

This gives us a more precise understanding of *when* children learn things and *in what order*

English past tense

Much productivity work discusses the English past tense, which has one "regular" pattern and a number of "irregular" patterns

	base	*past*	
'climb'	klajm	klajmd	
'stop'	stɑp	stɑpt	*suffix*
'melt'	mɛlt	mɛltəd	
'hit'	hɪt	hɪt	*no change*
'come'	kəm	kejm	*vowel change*
'sing'	sɪŋ	sæŋ	
'creep'	krijp	krɛpt	*vowel change + suffix*
'teach'	tijtʃ	tɔt	*vowel + consonant change*
'go'	gow	wɛnt	*suppletion (allomorphy)*

English past tense

Much productivity work discusses the English past tense, which has one "regular" pattern and a number of "irregular" patterns

	base	*past*	
'climb'	klajm	klajmd	
'stop'	stɑp	stɑpt	*suffix*
'melt'	mɛlt	mɛltəd	
'hit'	hɪt	hɪt	*no change*
'come'	kəm	kejm	*vowel change*
'sing'	sɪŋ	sæŋ	
'creep'	krijp	krɛpt	*vowel change + suffix*
'teach'	tijtʃ	tɔt	*vowel + consonant change*
'go'	gow	wɛnt	*suppletion (allomorphy)*

- "Irregulars": higher *token frequency* (85% of past forms heard by children)
- "Regulars": higher *type frequency* (86% of 1000 most frequent verbs)

Overregularization

Occasionally, verbs with irregular past tenses are given the regular ending: [gowd], [krijpt]

- Children go through a phase of overregularization
- Adults sometimes overregularize as speech errors
- The opposite ([lijk] → [lejk]) is very rare, though has happened historically ([dajvd] or [dowv])

Overregularization

Occasionally, verbs with irregular past tenses are given the regular ending: [gowd], [krijpt]

- Children go through a phase of overregularization
- Adults sometimes overregularize as speech errors
- The opposite ([lijk] → [lejk]) is very rare, though has happened historically ([dajvd] or [dowv])

In fact, children often follow a *U-shaped pattern* of learning irregular verbs (Marcus et al., 1992):

1. After an initial learning period, irregulars are marked correctly (if at all)
2. For a time, some irregulars are given the regular ending
3. Overregularization reduces and irregulars are marked correctly again

U-shaped learning curve

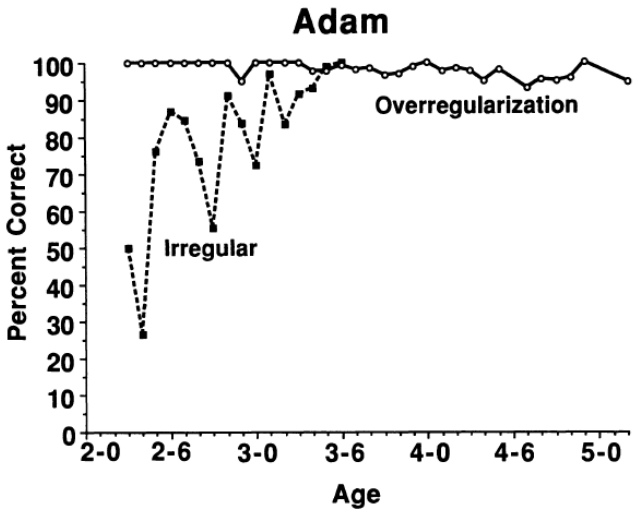


FIG. 33.—Proportion of Adam's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data but are shown in the curve at the top (subtracted from 100%).

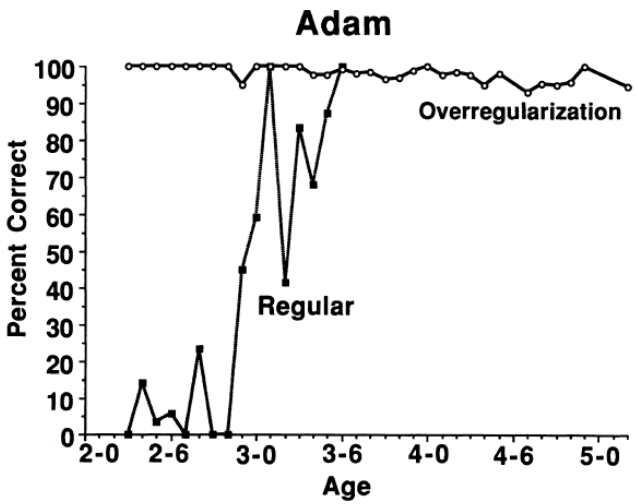


FIG. 34.—Proportion of Adam's regular verbs in obligatory past tense contexts that were correctly marked for tense and his overregularization rate (subtracted from 100%).

U-shaped learning curve

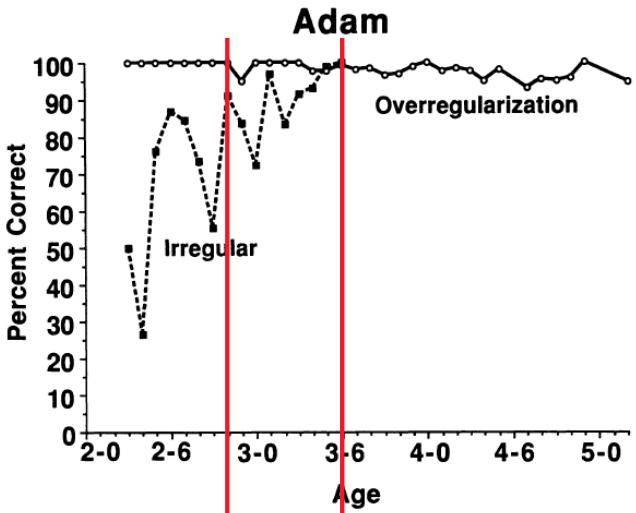


FIG. 33.—Proportion of Adam's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data but are shown in the curve at the top (subtracted from 100%).

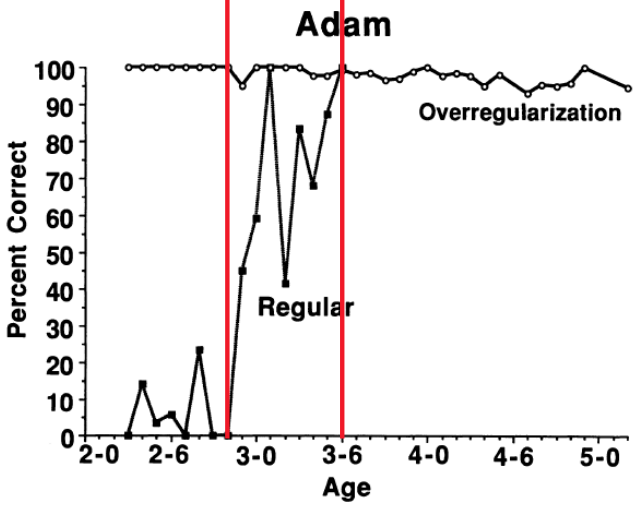


FIG. 34.—Proportion of Adam's regular verbs in obligatory past tense contexts that were correctly marked for tense and his overregularization rate (subtracted from 100%).

Phases of learning the past tense:

- Learning to mark irregulars, regulars often unmarked
- Learning to mark regulars, some overregularization
- Past tense consistently marked, more overregularization
- Regulars and irregulars formed correctly (not shown)

U-shaped learning curve

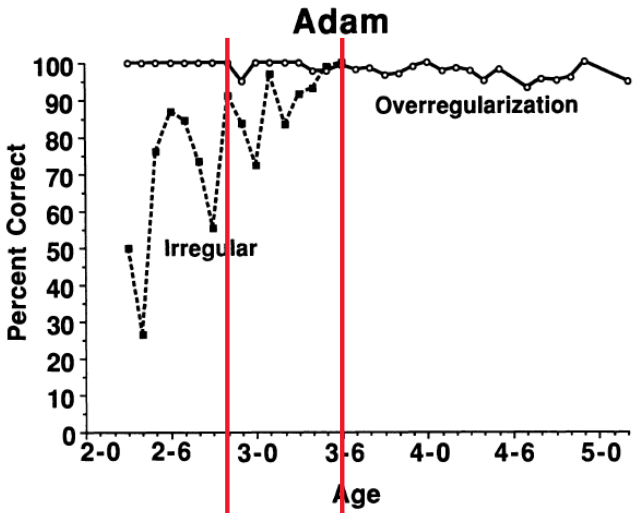


FIG. 33.—Proportion of Adam's irregular verbs in obligatory past tense contexts that were correctly marked for tense. Overregularizations do not enter into these data but are shown in the curve at the top (subtracted from 100%).

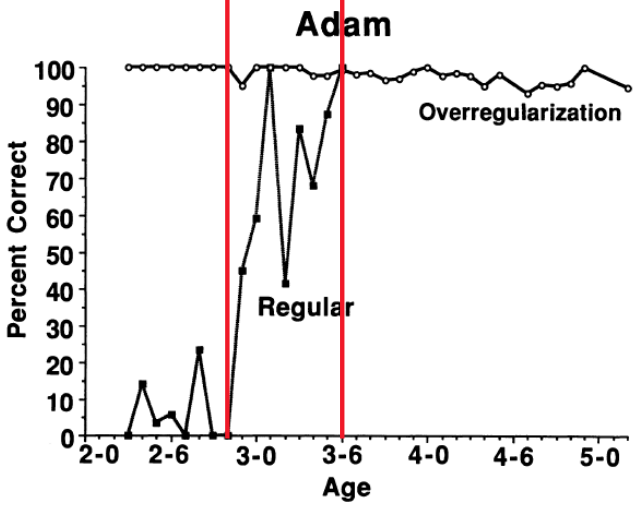


FIG. 34.—Proportion of Adam's regular verbs in obligatory past tense contexts that were correctly marked for tense and his overregularization rate (subtracted from 100%).

Phases of learning the past tense:

- Learning to mark irregulars, regulars often unmarked
- Learning to mark regulars, some overregularization
- Past tense consistently marked, more overregularization
- Regulars and irregulars formed correctly (not shown)

These correspond to learning stages:

1. Memorizing the past tense of each verb separately
2. Learning the regular [d] rule(s) and overapplying it
3. Sorting out the exceptions

Default behavior

Wug tests are just one example of default behavior where we'd expect regular inflection (Marcus et al., 1995, p. 197):

- onomatopeia: the bell [dɪŋd]
- quotations: I found three [mænz] on that page
- names: Julia and the rest of the [tʃajldz]
- truncations: we [sɪŋkt] up the sound with the video
- unassimilated borrowings: we ate [lɑŋgowʃəz]
- acronyms: he got [dijɛfejd]
- derived forms: she [gɹændstændəd], there have been several [bætmænz], a group of [lowlajfs], no [ɪfs] [ændz] or [bəts]

In English, this role is fulfilled by [z s əz] for the plural and [d t ət] for the past tense (depending on phonology)

Default behavior

Wug tests are just one example of default behavior where we'd expect regular inflection (Marcus et al., 1995, p. 197):

- onomatopeia: the bell [dɪŋd]
- quotations: I found three [mænz] on that page
- names: Julia and the rest of the [tʃajldz]
- truncations: we [sɪŋkt] up the sound with the video
- unassimilated borrowings: we ate [lɑŋgowʃəz]
- acronyms: he got [dijɛfejd]
- derived forms: she [gɹændstændəd], there have been several [bætmænz], a group of [lowlajfs], no [ɪfs] [ændz] or [bəts]

In English, this role is fulfilled by [z s əz] for the plural and [d t ət] for the past tense (depending on phonology) – what about your language?

German plurals

German has five plural suffixes that sometimes cooccur with vowel alternation (umlaut) (McCurdy et al., 2020, p. 1746):

suffix	*example*	*singular*	*plural*	*types*	*tokens*
(ə)n	‘street’	ʃtʁaːsə	ʃtʁaːsən	48%	45%
ə	‘dog’	hʊnt	hʊndə	27%	21%
	‘cow’	kuː	kyːə		
Ø	‘thumb’	daʊmən	daʊmən	17%	29%
	‘mother’	mʊtɐ	mʏtɐ		
ɐ	‘child’	kɪnt	kɪndɐ	4%	3%
	‘forest’	valt	vɛldɐ		
s	‘car’	aʊto	aʊtos	4%	2%

How do German words get their plurals?

APPENDIX 2:
A SUMMARY OF THE GERMAN PLURAL SYSTEM (BASED ON MUGDAN, 1977)

List 1 → *-e* + change stem as indicated.

List 2 → Change stem and add suffix as indicated

Coda = ∂:
- Neuter with prefix *ge-* → -0.
- other neuter → *-en*.
- *Käse* → -0.
- Other Masculine or Feminine → *-en*.

Coda = non-∂ vowel:
- Feminine with coda = long stressed /e/ or /i/ → *-en*
- other Feminine → *-s*
- Masculine or Neuter → *-s*.

Coda = ∂ + consonant:
- Feminine *Mutter, Tochter* → 0 + umlaut
- other Feminine → *-en*
- *Kloster* → 0 + umlaut
- other Neuter → 0
- Masculine in List 3 → *-en*
- Masculine in List 4 → 0 + umlaut
- other Masculine → 0

Other Consonant-final Codas:
- Feminine in List 5 → *-s*
- Feminine with *-sal* suffix → *-e*
- Feminine in List 6 → *-e* + umlaut
- Other Feminine → *-en*
- Neuter with *-lein* suffix → 0
- Neuter *Floß* → *-e* + umlaut
- Neuter in List 7 → *-s*
- Neuter in List 8 → *-en*
- Neuter in List 9 → *-er* + umlaut
- other Neuter → *-e*
- Masculine in List 10 → *-en*
- Masculine in List 11 → *-s*
- Masculine in List 12 → *-er* + umlaut
- Masculine animate in List 13 → *-en*
- Masculine animate in List 15 → *-e* + umlaut
- other Masculine animate → *-e*
- Masculine inanimate in List 14 → *-en*
- other Masculine inanimate → *-e*

List 1: *Character-Charactere, Klima-Klimate, Lira-Lire*

List 2: *Numerus-Numeri, Jus-Jura, Fortis-Fortes, Koma-Komata, Cherub-Cherubim*

List 3: *Brösel, Bauer, Bayer, Fussel, Friesel, Gevatter, Hummer, Kaffer, Lackel, Maskel, Model, Pantoffel, Pommer* (nonstandard) *Stachel, Stiefel, Stummel, Stoppel, Treidel, Vetter, Ziegel*

List 4: *Acker, Apfel, Boden, Bogen, Bruder, Faden, Garten, Graben, Hafen, Hammel, Hammer, Handel, Kasten, Kragen, Laden, Magen, Mangel, Mantel, Nabel, Nagel, Ofen, Polster, Rodel, Sattel, Schaden, Schnabel, Schochen, Schwager, Stadel, Stofel, Vater, Vogel, Wagen*

List 5: *Abming, Bar, Brigg, Creme, Flak, Gig, Hall, Kautsch, Lok, Mansell, Ordre, Pak, Persenning, Pipeline, Trum*

List 6: *Angst, Bank* (bench), *Braut, Brunft, Brunst, Faust, -flucht* (Aus-), *Frucht, Gans, Gruft, Hand, Haut, Kraft, -kunft, Kluft* (cleft), *Kuh, Kunst, Laus, Luft, Lust, Macht,*

Magd, Maus, Nacht, Naht, Niß, Not, Nuß, Sau, Schlucht, Schnur, Schwulst, Stadt, Sucht, Wand, Wulst, Wurst, -zucht (Ab-), *Zunft, -uts*

List 7: *Akkordeon, Barsten, Bilett, Bisam, Biwak, Brikett, Büfett, Camp, Davit, Deck, Dessert, Dock, Ensemble, Fort, Gatt, Gig, Girl, Haff, Handikap, Heck, Helikon, Hoch, Hotel, Interim, Jackett, Jam, Kabarett, Kaff, Kajak, Kap, Klosett, Kolon, Kornett, Korpa, Korsett, Kotelett, Kuvert, Labor, Meeting, Myom, Pickett, Picknick, Pingpong, Plaid, Part, Regal* (Satz-; set of shelves), *Regime, Relief, Repertoire, Requiem, Schrapnell, Shalom, Slum, Soll, Souvenir, Spill, Spiritual, Steak, Tamtam, Tandem, Tarock, Team, Tief, Treff* (club in playing cards), *Triptyk, Tunnel, Wrack*

List 8: *Apokryph, Autograph, Bett, Blag, Chromasom, Eck, Gatt, Gör, Hemd, Herz, Insekt, Ion, Juwel, Leik, Ohr, Quant, Requisit, Spant, Stag, Statut, Verb, Wrat*

List 9: *Aas, Amt, Bad, Balg, Band, Bein, Biest, Bild, Blatt, Brett, Brot, Buch, Dach, Daus, Ding, Dorf, Ei, Fach, Faß, Feld, Gehalt, Geld, Gemach, Gemüt, Geschlecht, Gesicht, Gespenst, Gewand, Glas, Glied, Grab, Gras, Gut, Haupt, Haus, Holz, Horn, Hospital, Huhn, Kalb, Kind, Kleid, Korn, Kraut, Lamm, Land, Licht, Lid, Lied, Loch, Mahl, Mal, Maul, Mensch, Moos* (swamp), *Nest, Pfand, Rad, Regiment, Reis, Rind, Roß, Scheit, Scheusal, Schild, Schloß, Schwert, Spital/Hospital, Stift, Stück, Tal, Trumm, Tuch, Viech, Volk, Wams, Weib, Weif, Wort, -tum*

List 10: *Augur, Herr* (spoken German), *Konsil, Modul, Mogul, Nachbar, Satyr, Triumvir, Ungar*

List 11: *Anorak, Arrak, Back, Baobab, Baribal, Bisam, Bison, Blizzard, Block, Bluff, Bob, Boulevard, Browning, Bulldog, Camembert, Champion, Chef, Clan, Clip, Clown, Code, Cut, Dekor, Dollar, Drain, Drink, Fasching, Flip, Flirt, Foulard, Frack, Galopp, Gnaur, Gong, Grog, Groom, Harem, Havelock, Jak, Job, Karneval, Kanon, Kerl, Klub, Kognak, Könsam, Kornett, Krepp, Leatnant, Lift, Lord, Marmgal, Mop, Muffion, Nabob, Park, Pemmikan, Portulak, Pudding, Puff* (brothel), *Pump* (shoe), *Radar, Raglan, Safe, Schal, Scheck, Scheich, Schock, Seal, Shag, Sheriff, Skunk, Smoking, Snob, Sowjet, Spurt, Star, Start, Steward, Stopp, Tank, Tarock, Tattersall, Test, Tic, Tip, Toast, Toboggan, Totem, Tramp, Trick, Trip, Truck, Tripp, Trust, Tschibuk, Tusch, Vamp, Yard*

List 12: *Balg, Daus, Dorn, Geist, Geschmack, Gott, Handgott, Klotz, Leib, Mann, Mund, Ort, Rand, Rest, Schoß* (Zoll), *Ski, Stein, Strauch, Trumm, Wald, Wanst, Wicht, Wiking, Warm, -tum* (Irrtum, Reichtum)

List 13: *Adjunkt, Ahn, Alk, Anthroposoph, Architekt, Astronaut, Autodidakt, Bär, Barbar, Beugstaß, Bojar, Bosniak, Bub, Bur, Bursch, Depp, Diakon, Dolmetsch, Falk, Fasan, Fex, Fink, Fratz, Fürst, Gams, Gast* (crewman), *Gaur, Geck, Gendarm, Genuß, Geodät, Gepard, Gesell, Gnom, Graf, Greif, Gymnasiast, Hagestolz, Heiduck, Held, Herr, Hintersaß, Hirt, Husar, Kadett, Kalif, Kalmück, Kamerad, Korsar, Kosak, Kosmonaut, Lakai, Lemur, Leopard, Lump, Maat, Magyar, Mameluck, Mann* (fiefs), *Mensch, Misanthrop, Mohr, Nachfahr, Narr, Oberst, Ochs, Pandur, Papagei, Partisan, Pfaff, Pfau, Phantast, Philanthrop, Philosoph, Polyp, Präfekt, Prinz, Proselyt, Protz, Rebell, Rekrut, Rival, Satrap, Schelm, Schenk, Scholar, Schultheiß, Schütz, Senn, Spatz, Spiritual, Starost, Steinmetz, Substitut, Tatar, Titan, Tor, Tribun, Truchseß, Tyrann, Ulan, Untertan, Vagabund, Vasall, Veteran, Vorfahr, Zar, Zentaur, Zyklop, -ând, -ânt, -arch, -ât, -ét, -ént, -éut, -gráph, -ik, -ist, -it, -ith, -ög, -ôm, -ôt, -ürg*

List 14: *Architrav, Arithmograph, Asch* (fish), *Autoklav, Automat, Basilisk, Dorn, Exponent, Fels, Firn, Gau, Hahn* (faucet), *Juwel, Logograph, Lorbeer, Mast, Molar, Mond* (month), *Nerv, Obelisk, Pakt, Psalm, Pulp, Satellit, Schmerz, Schoß* (Zoll), *Schwad, See, Sproß, Staat, Strahl, Typ, Zeh, Zink, Zins, -ând, -ânt, -ênd, -ât, -gráph, -iént, -phyt*

List 15: *Abt, Admiral, Altar, Anfang, Anwalt, Arsch, Arzt, Asch, Ast, Bach, Balg, Ball, Band, Bart, Baß, Bauch, Baum, Bausch, Belag, Bischof, Block, Bock, Brand, Brauch, Bruch, Bug, Bult, Bund, Busch, Chor, Choral, Damm, Dampf, Darmn, Diebstahl, Drahi, Drang, Druck* (generally), *Duft, Dunst, Fall, Fang, Floh, Fluch, Flor, Flug, Fluß, Frack, Frosch, Frost, Fuchs, Furz, Fuß, Gang, Gast, Gauch, Gaul, General, Genuß, Geschmack, Graf, Grund, Gruß, Guß, Hahn, Hals, Hang, Herzog, Hof, Hub, Hut, Kahn, Kamm, Kampf, Kanal, Kaplan, Kardinal, Kauf, Kauz, Klang, Klaps, Klops, Kloß, Klotz, Knall, Knauf, Knaul, Knopf, Knuff, Knust, Koch, Kopf, Korb, Korporal, Krach, Krampf, Kran, Kranz, Kropf, Krug, Kuß, -laß, Latz, Lauf, Lohn, Markt, Marsch, Marschall, Matz, Mops, Morast, Mund, Napf, Pack, Palast, Papst, Pasch, Paß, Pastör, Pfahl, Pflock, Pflug, Plan, Platz, Pott, Prahm, Propf, Propst, Puff* (nudge, dig), *Punsch, Rang, Raun, Rausch, Rock, Ruch, Rumpf, Saal, Sack, Saft, Sand, Sang, Sarg, Satz, Saum, Schacht, Schaft, Schalk, Schall, Schank, Schatz, Schaum, Schlag, Schlamm, Schlauch, Schlot, Schluck, Schlund, Schlupf, Schluß, Schmatz, Schamus, Schmack, Schnaps, Schopf, Schoß, Schram, Schrank, Schub, Schurf, Schurz, Schuß, Schwamm, Schwan, Schwank, Schwanz, Schwarm, Schwulst, Schwung, Schwur, Sohn, Span, spaß, Spat, Spruch, Sprung, Spund, Stab, Stahl, Stall, Stamm, Stand, Staubm, Stock, Storch, Stoß, Strang, Strauß* (flowers), *Strom, Stumpf, Stuhl, Stumpf, Sturm, Sturz, Sumpf, Tanz, Tenör, Ton* (tone), *Topf, Torf, -trag, Tram, Trunk, Traum, Trog, Topf, Trumpf, Trunk, Turm, Verdruß, Vogt, Wall, Wams, -wand, Wanst, Wolf, Worb, Wuchs, Wulst, Wunsch, Wurf, Zahn, Zaum, Zaun, Zoll, Zug, Zwang, Zwieback*

Infrequent defaults

Marcus et al. (1995): In German, this role for plural nouns is fulfilled by [s], the least frequent suffix

- onomatopeia: [vaʊvaʊs] ‘bow-wows’
- quotations: [mans] ‘“man”s’
- names: Thomas [mans]
- truncations: [vɛsis] “Westies”
- unassimilated borrowings: [kiːɔsks]
- acronyms: [beːɛmveː] ‘BMWs’
- derived forms: [aːbɐs] ‘ors’, [faʊsts] ‘Fausts’, [gɔlfs] ‘Golfs’

When these end in sibilants, the more common [ə] and [(ə)n] are used: [bɔsə] ‘bosses’, [mɛʁtseːdəsə] ‘Mercedeses’, [matsən] ‘magnetic recordings’ (Indefrey, 1999)

German wug tests

Adults and children use [s] and the two most common [ə] and [(ə)n] (Zaretsky & Lange, 2016, p. 165):

plural		*% rhymes*	*% non-rhymes*
	umlaut	0.2	0.04
ɐ	+ umlaut	6	2
ɐ		8	4
ən	(+ umlaut)	20	27
ə	+ umlaut	14	6
ə		44	46
s	(+ umlaut)	8	15

- [s] is more common with "non-rhymes" – phonologically legal but deviant words that don't rhyme with any existing words in German like [fnøːk]
- Other suffixes have preferences as well, like [ən] for feminines
- While [s] is the typical default "emergency" suffix, it's not the only productive one

Lack of regularization

Two genitive suffixes, [a] and [u], for Polish masculine inanimate nouns (Dąbrowska, 2001, 2008):

- [twɔka] ‘piston’ vs. [twɔku] ‘crowd’
- No consistent generalization, phonological or otherwise
- [a] is more common (70–80% of types and tokens)

Lack of regularization

Two genitive suffixes, [a] and [u], for Polish masculine inanimate nouns (Dąbrowska, 2001, 2008):

- [twɔka] ‘piston’ vs. [twɔku] ‘crowd’
- No consistent generalization, phonological or otherwise
- [a] is more common (70–80% of types and tokens)

Both are used in various “default” situations:

- truncations: [mɛrtsa] ‘Mercedes’, [haszu] ‘hash’
- borrowings: [fɔnɛmu] ‘phoneme’, [driŋka] ‘drink’
- acronyms: [panu] ‘PAN (Polish Academy of Sciences’
- derived forms: [sʂɨvatʂa] ‘stapler’, [zbʲoru] ‘collection’, [kɔntomʲɛzɻa] ‘protractor’, [trujzɛmbu] ‘trident’

Neither is really “regular” or even “default”

Lack of regularization

Two genitive suffixes, [a] and [u], for Polish masculine inanimate nouns (Dąbrowska, 2001, 2008):

- [twɔka] 'piston' vs. [twɔku] 'crowd'
- No consistent generalization, phonological or otherwise
- [a] is more common (70–80% of types and tokens)

Both are used in various "default" situations:

- truncations: [mɛrtsa] 'Mercedes', [haszu] 'hash'
- borrowings: [fɔnɛmu] 'phoneme', [driŋka] 'drink'
- acronyms: [panu] 'PAN (Polish Academy of Sciences'
- derived forms: [sʂɨvatʂa] 'stapler', [zbʲoru] 'collection', [kɔntomʲɛzɻa] 'protractor', [trujzɛmbu] 'trident'

Neither is really "regular" or even "default"

- Polish children start to use the genitive very early
- Way fewer errors than English children make in the past tense
- More [u] → [a] errors than [a] → [u] errors, but much less lopsided than in English

Summary

- Children sometimes overgeneralize very common word formation patterns
- We can identify cases that expect a default, where memory (and other possibilities) fail
- Overgeneralization, default behavior, and frequency often align, but not always
- Sometimes we can't identify any pattern as a default or regular at all

Outline

Modelling productivity

Many different ideas for how to account for productivity (acquisition, wug tests, etc.)

- formal abstract rules vs. analogical comparison
- categorical vs. probabilistic application
- regular phonology/morphology vs. separate pattern-matching component

Structured analogy

So far we've looked at the spread of English past /d/ – now let's look at another class (Bybee & Moder, 1983)

	present	*past*
'spin'	spɪn	spən
'win'	wɪn	wən
'cling'	clɪŋ	cləŋ
'fling'	flɪŋ	fləŋ
'sling'	slɪŋ	sləŋ
'sting'	stɪŋ	stəŋ
'string'	stɹɪŋ	stɹəŋ
'swing'	swɪŋ	swəŋ
'wring'	ɹɪŋ	ɹəŋ
'hang'	hæŋ	həŋ
'slink'	slɪŋk	sləŋk
'stick'	stɪk	stək
'strike'	stɹajk	stɹək
'sneak'	snijk	snək
'dig'	dɪg	dəg

Structured analogy

So far we've looked at the spread of English past /d/ – now let's look at another class (Bybee & Moder, 1983)

	present	*past*
'spin'	spɪn	spən
'win'	wɪn	wən
'cling'	clɪŋ	cləŋ
'fling'	flɪŋ	fləŋ
'sling'	slɪŋ	sləŋ
'sting'	stɪŋ	stəŋ
'string'	stɹɪŋ	stɹəŋ
'swing'	swɪŋ	swəŋ
'wring'	ɹɪŋ	ɹəŋ
'hang'	hæŋ	həŋ
'slink'	slɪŋk	sləŋk
'stick'	stɪk	stək
'strike'	stɹajk	stɹək
'sneak'	snijk	snək
'dig'	dɪg	dəg

These verbs tend to share several properties that together make a *prototype* for the class:

Structured analogy

So far we've looked at the spread of English past /d/ – now let's look at another class (Bybee & Moder, 1983)

	present	*past*
'spin'	spɪn	spən
'win'	wɪn	wən
'cling'	clɪŋ	cləŋ
'fling'	flɪŋ	fləŋ
'sling'	slɪŋ	sləŋ
'sting'	stɪŋ	stəŋ
'string'	stɹɪŋ	stɹəŋ
'swing'	swɪŋ	swəŋ
'wring'	ɹɪŋ	ɹəŋ
'hang'	hæŋ	həŋ
'slink'	slɪŋk	sləŋk
'stick'	stɪk	stək
'strike'	stɹajk	stɹək
'sneak'	snijk	snək
'dig'	dɪg	dəg

These verbs tend to share several properties that together make a *prototype* for the class:

- [ŋ] at the end

Structured analogy

So far we've looked at the spread of English past /d/ – now let's look at another class (Bybee & Moder, 1983)

	present	*past*
'spin'	spɪn	spən
'win'	wɪn	wən
'cling'	clɪŋ	cləŋ
'fling'	flɪŋ	fləŋ
'sling'	slɪŋ	sləŋ
'sting'	stɪŋ	stəŋ
'string'	stɹɪŋ	stɹəŋ
'swing'	swɪŋ	swəŋ
'wring'	ɹɪŋ	ɹəŋ
'hang'	hæŋ	həŋ
'slink'	slɪŋk	sləŋk
'stick'	stɪk	stək
'strike'	stɹajk	stɹək
'sneak'	snijk	snək
'dig'	dɪg	dəg

These verbs tend to share several properties that together make a *prototype* for the class:

- [ŋ] at the end
- [sC] at the start

Structured analogy

So far we've looked at the spread of English past /d/ – now let's look at another class (Bybee & Moder, 1983)

	present	*past*
'spin'	spɪn	spən
'win'	wɪn	wən
'cling'	clɪŋ	clən
'fling'	flɪŋ	flən
'sling'	slɪŋ	slən
'sting'	stɪŋ	stən
'string'	stɹɪŋ	stɹən
'swing'	swɪŋ	swən
'wring'	ɹɪŋ	ɹən
'hang'	hæŋ	hən
'slink'	slɪŋk	slənk
'stick'	stɪk	stək
'strike'	stɹajk	stɹək
'sneak'	snijk	snək
'dig'	dɪg	dəg

These verbs tend to share several properties that together make a *prototype* for the class:

- [ŋ] at the end
- [sC] at the start
- [ɪ] in the middle

Structured analogy

So far we've looked at the spread of English past /d/ – now let's look at another class (Bybee & Moder, 1983)

	present	*past*
'spin'	spɪn	spən
'win'	wɪn	wən
'cling'	clɪŋ	clən
'fling'	flɪŋ	flən
'sling'	slɪŋ	slən
'sting'	stɪŋ	stən
'string'	stɹɪŋ	stɹən
'swing'	swɪŋ	swən
'wring'	ɹɪŋ	ɹən
'hang'	hæŋ	hən
'slink'	slɪŋk	slənk
'stick'	stɪk	stək
'strike'	stɹajk	stɹək
'sneak'	snijk	snək
'dig'	dɪg	dəg

These verbs tend to share several properties that together make a *prototype* for the class:

- [ŋ] at the end
- [sC] at the start
- [ɪ] in the middle

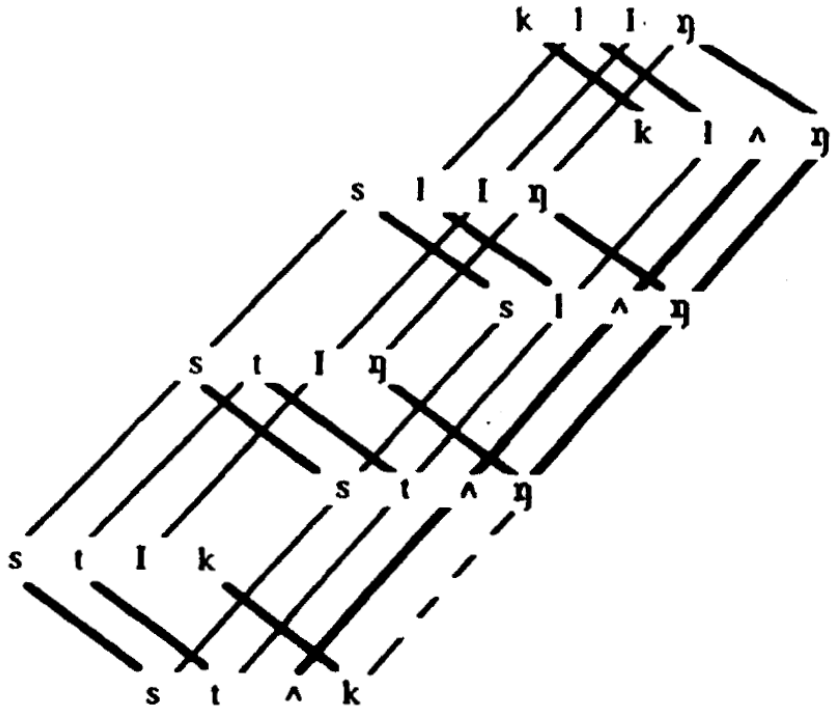
Wugs closer to the prototype were given past [ə] more often

shape	% [ə]	*shape*	% [ə]
sCCɪŋ(k)	44%	sCCɪŋ(k)	44%
sCɪŋ(k)	37%	sCCɪ{k/g}	25%
CCɪŋ(k)	27%	sCCɪ{m/n}	21%
Cɪŋ(k)	22%	sCCɪC	4%

Analogical models

This "prototype effect" can be explained through *analogy* (e.g. Bybee, 1995):

- Speakers store (frequent) inflected forms
- When they need to make a new form of a word, they compare it to similar existing forms
- The more similar the forms are to the new word (and each other), the more influence they will have on the new forms



Source- and product-oriented generalizations

Analogical models often capture two types of generalizations:

- *source-oriented*: present in the base
- *product-oriented*: present in the inflected form

	present	*past*
'string'	stɹɪŋ	stɹəŋ

Source- and product-oriented generalizations

Analogical models often capture two types of generalizations:

- *source-oriented*: present in the base (sCC, ɪ, ŋ)
- *product-oriented*: present in the inflected form (sCC, ə, ŋ)

	present	*past*
'string'	stɹɪŋ	stɹəŋ

Source- and product-oriented generalizations

Analogical models often capture two types of generalizations:

- *source-oriented*: present in the base (sCC, ɪ, ŋ)
- *product-oriented*: present in the inflected form (sCC, ə, ŋ)

	present	*past*
'string'	stɹɪŋ	stɹəŋ

A product-oriented generalization: most English past tense verbs end in alveolar stops [d t]

Source- and product-oriented generalizations

Analogical models often capture two types of generalizations:

- *source-oriented*: present in the base (sCC, ɪ, ŋ)
- *product-oriented*: present in the inflected form (sCC, ə, ŋ)

	present	*past*
'string'	stɹɪŋ	stɹəŋ

A product-oriented generalization: most English past tense verbs end in alveolar stops [d t]

	present	*past*
'put'	pʊt	pʊt
'set'	sɛt	sɛt
'quit'	kwɪt	kwɪt
'spread'	spɹɛd	spɹɛd

Product-oriented generalizations in Hausa

In Hausa, plurals are formed *entirely* by product-oriented generalizations, including in nonce word studies (Bybee, 1995, p. 445):

	singular	*plural*
‘gown’	rìigáa	ríigúnàa
‘store’	kàntíi	kántúnàa
‘porch’	záurèe	záurúkàa
‘hand’	hán:úu	hán:úwàa
‘prosperity’	árzìkíi	árzúkàa
‘sword’	tákòobii	tákúbàa

These patterns apply to bases with particular phonological shapes, but can be formed in different ways.

Product-oriented generalizations in Hausa

In Hausa, plurals are formed *entirely* by product-oriented generalizations, including in nonce word studies (Bybee, 1995, p. 445):

	singular	*plural*
‘gown’	rìigáa	ríigúnàa
‘store’	kàntíi	kántúnàa
‘porch’	záurèe	záurúkàa
‘hand’	hán:úu	hán:úwàa
‘prosperity’	árzìkíi	árzúkàa
‘sword’	tákòobii	tákúbàa

These patterns apply to bases with particular phonological shapes, but can be formed in different ways.

Regulars by analogy

When we see a “regular” pattern (like the English past tense), this is just a particularly powerful target of analogy

- Widest range of forms to which it applies
- High type frequency (but not always!)

Analogical models sometimes fail to choose the right “regular” past tense (Albright & Hayes, 2003, p. 151):

	singular	*plural?*	*bad influences*
‘render’	ɹɛnd̩	ɹɛnd̩əd	ɹɛnd, ɛnd, ɹɛnt, vɛnd, ɹejd, fɛnd, mɛnd, tɛnd, ɹawnd, dɹɛd
‘whisper’	wɪspɹ̩	wɪspɹ̩t	wɪp, wɪʃ, wɪsk, wɪns, kwɪp, lɪsp, swɪʃ, ɹɪp, wɹ̩k, mɪs

Analogical models: an example

To actually make an analogical model, we need to be explicit about how to compare forms to one another

One way of doing this: align segments and assess featural similarity (Albright & Hayes, 2003, p. 131)

shine:	ʃ		null		null		a		ɪ		n	
penalty:	0.155	+	0.6	+	0.6	+	0	+	0	+	0.667	= 2.022
scride:	s		k		r		a		ɪ		d	

Then find overall similarity scores (transformed) for all patterns:

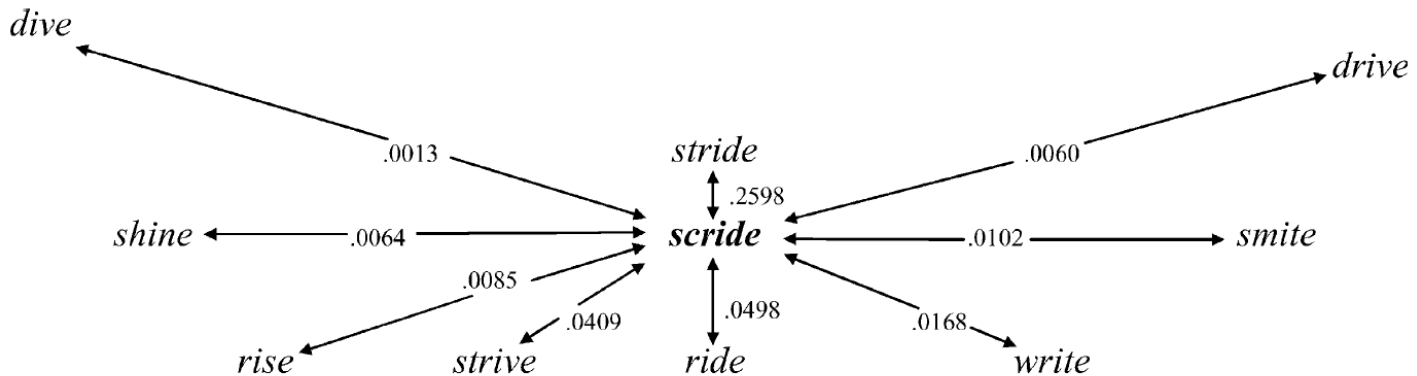
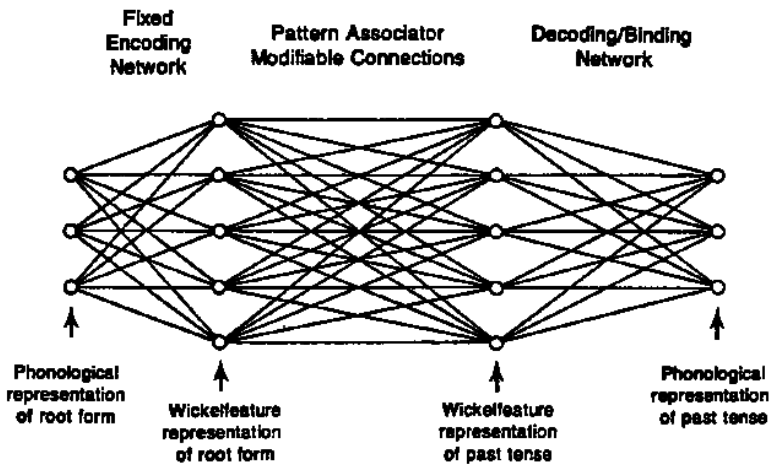


Fig. 1. Similarity of all [aɪ] → [o] forms to *scride*.

QUESTION: Is this source-oriented, product-oriented, or both?

Neural networks

Rumelhart and McClelland (1986): You don't need to store past tense forms, they can arise from weights in a neural network

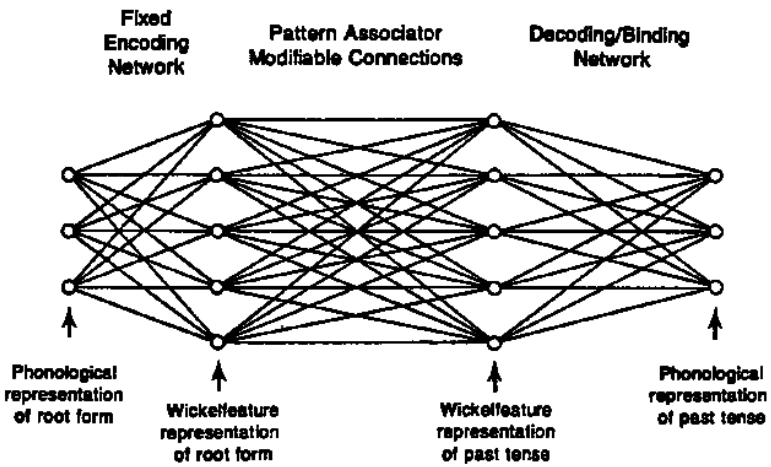


Words are represented as triplets of phonological features

- [kəm]: #-k-ə / k-ə-m / ə-m-#
- [k]: occlusive, stop, back, unvoiced
- [ə]: vowel, high, middle, short
- [ej]: vowel, low, front, long
- [m]: occlusive, nasal, front, voiced
- [kəm]: … / stop-vowel-nasal / unvoiced-middle-voiced / …

Neural networks

Rumelhart and McClelland (1986): You don't need to store past tense forms, they can arise from weights in a neural network



Words are represented as triplets of phonological features

- [kəm]: #-k-ə / k-ə-m / ə-m-#
- [k]: occlusive, stop, back, unvoiced
- [ə]: vowel, high, middle, short
- [ej]: vowel, low, front, long
- [m]: occlusive, nasal, front, voiced
- [kəm]: … / stop-vowel-nasal / unvoiced-middle-voiced / …

Task: convert present Wickelfeatures to past Wickelfeatures:

- unvoiced-middle-voiced → unvoiced-front-voiced

The problem with the neural network

This neural network had similar problems to the analogical models described previously (Pinker & Prince, 1988)

- Strange, unrealistic errors: [mejl] → [mɛmb̩ld], [skwɔt] → [skwɔkt]
- Double marking: [stɛp] → [stɛptəd]
- It can only account for overregularization under the assumption that children first see large amounts of irregulars, then see large amounts of regulars – but this isn't what happens (about 50% of their types are consistently regular)
- (many more things, their paper is 121 pages)

The problem with the neural network

This neural network had similar problems to the analogical models described previously (Pinker & Prince, 1988)

- Strange, unrealistic errors: [mejl] → [mɛmb̩ld], [skwɔt] → [skwɔkt]
- Double marking: [stɛp] → [stɛptəd]
- It can only account for overregularization under the assumption that children first see large amounts of irregulars, then see large amounts of regulars – but this isn't what happens (about 50% of their types are consistently regular)
- (many more things, their paper is 121 pages)

Pinker and Prince (1988) argue that the problems are inherent to any neural network

The problem with the neural network

This neural network had similar problems to the analogical models described previously (Pinker & Prince, 1988)

- Strange, unrealistic errors: [mejl] → [mɛmb̩ld], [skwɔt] → [skwɔkt]
- Double marking: [stɛp] → [stɛptəd]
- It can only account for overregularization under the assumption that children first see large amounts of irregulars, then see large amounts of regulars – but this isn't what happens (about 50% of their types are consistently regular)
- (many more things, their paper is 121 pages)

Pinker and Prince (1988) argue that the problems are inherent to any neural network – but things have changed since 1986 (LLMs)

Contemporary neural networks

Kirov and Cotterell (2018): A contemporary neural network with more flexible representation (no Wickelfeatures) does much better on the English past tense

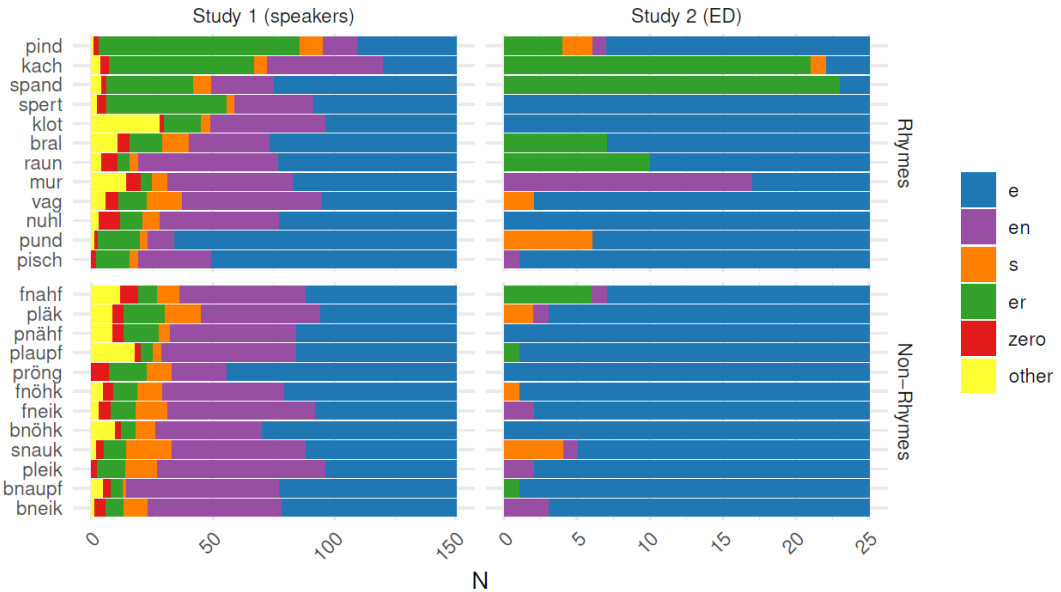
- Correctly learns almost all regular and irregular verbs
- Most errors on new words are overregularization errors
- No U-shaped curve in “acquisition”

Contemporary neural networks

Kirov and Cotterell (2018): A contemporary neural network with more flexible representation (no Wickelfeatures) does much better on the English past tense

- Correctly learns almost all regular and irregular verbs
- Most errors on new words are overregularization errors
- No U-shaped curve in “acquisition”

McCurdy et al. (2020): A similar neural network *does not* show human-like performance on German plural wug test



Dual route models

Pinker and Prince (1988): a strong division between *regular* and *irregular* morphology

- A single rule (add /d/) makes most English past-tense forms: /lijk+d/ 'leak'
- All others are stored in full: /spowk/ 'speak'

Evidence presented for this:

- Default use of /d/ (onomatopoeia, derived forms, etc.)
- U-shaped curve as period of learning the /d/ rule and the stored forms
- Lexical access: access speed of irregulars correlated with frequency, regulars are slower and inhibited by irregular "distractors"
- The pattern of use of /d/ in wug tests

Dual route models

Pinker and Prince (1988): a strong division between *regular* and *irregular* morphology

- A single rule (add /d/) makes most English past-tense forms: /lijk+d/ 'leak'
- All others are stored in full: /spowk/ 'speak'

Evidence presented for this:

- Default use of /d/ (onomatopoeia, derived forms, etc.)
- U-shaped curve as period of learning the /d/ rule and the stored forms
- **Lexical access: access speed of irregulars correlated with frequency, regulars are slower and inhibited by irregular "distractors"**
- **The pattern of use of /d/ in wug tests**

Lexical access

A common task in psycholinguistics: *lexical decision* (deciding whether a string of text is a valid word)

- Option 1: Find the word in storage (more frequent words are accessed more quickly)
- Option 2: Compose the word from stored parts (more frequent parts are accessed more quickly, plus a fixed processing time)

Dual route models predict: regulars (can) have effect of *stem frequency*, irregulars must have effect of *word frequency* (Pinker & Ullman, 2002)

- Some studies show support for this division (e.g. Yang, 2016)
- Others show more evidence of storage of regulars and more complicated effects of inflectional paradigms more generally (e.g. Milin et al., 2009)

Weird regulars

Prasada and Pinker (1993): If regulars and irregulars are both formed by analogy, both should show "neighborhood effects"

- Pseudo-irregular: prototypical [splɪŋ], intermedate [nɪŋ], distant [nɪst]
- Pseudo-regular: prototypical [plɪp], intermediate [smejg], distant [plowmf]

Weird regulars

Prasada and Pinker (1993): If regulars and irregulars are both formed by analogy, both should show "neighborhood effects"

- Pseudo-irregular: prototypical [splɪŋ], intermedate [nɪŋ], distant [nɪst]
- Pseudo-regular: prototypical [plɪp], intermediate [smejg], distant [plowmf]

In fact, participants used /d/ even for distant pseudo-regulars (they were allowed to supply multiple past tense forms

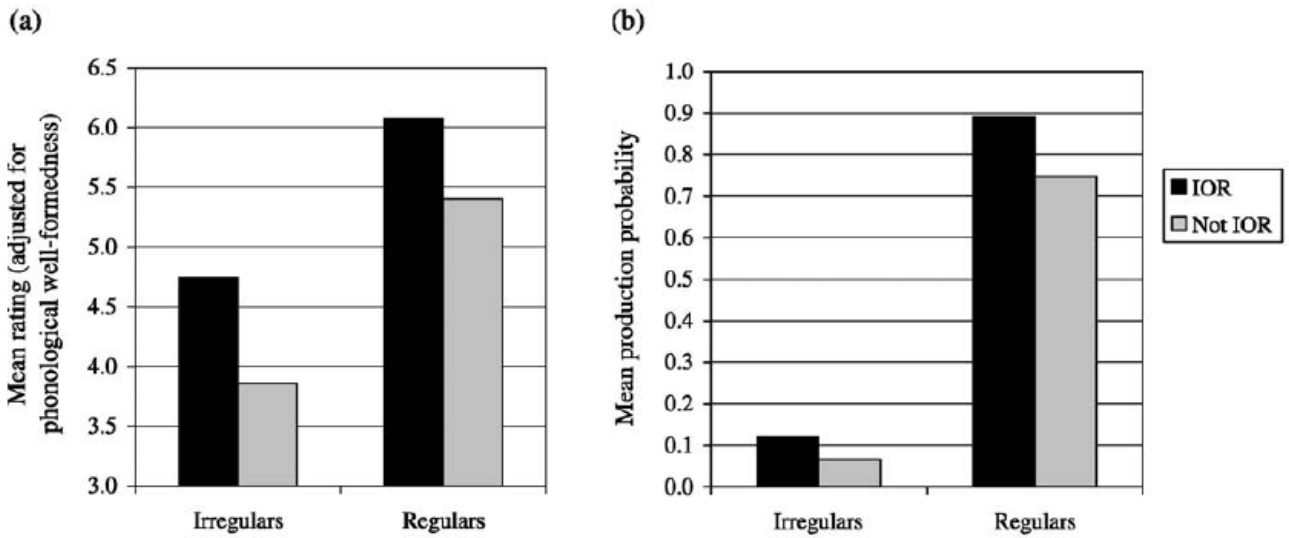
	pseudo-irregular		*pseudo-regular*	
	% non-suffixed	*rating*	*% suffixed*	*rating*
prototypical	59%	4.7	95%	5.5
intermediate	46%	4.2	92%	5.2
distant	37%	4.2	89%	5.3

Some analogical models and neural networks (e.g. Rumelhart & McClelland, 1986) do worse on these regulars

Islands of reliability

Albright and Hayes (2003): The weird regulars in Prasada and Pinker (1993) were too weird

- Island of reliability for regulars: [bɹɛdʒ], [gɛz], [nejs]
- Island of reliability for irregulars: [flijp], [glijd], [splɪŋ]
- Island of reliability for both: [dajz], [fɹow], [ɹajf]
- Island of reliability for neither: [guwd], [nəŋ], [prijk]



Unexpected if regulars are formed with a single rule!

Minimal generalization

Albright and Hayes (2003): There are *many rules* for the past tense, which we derive by combining more specific rules into more general ones

- Ø → əd / [vowt___]$_{[+past]}$
- Ø → əd / [nijd___]$_{[+past]}$

Minimal generalization

Albright and Hayes (2003): There are *many rules* for the past tense, which we derive by combining more specific rules into more general ones

- Ø → əd / [vowt___]$_{[+past]}$
- Ø → əd / [nijd___]$_{[+past]}$
- Ø → əd / [X [+cor, +ant, −nas, −cont]___]$_{[+past]}$

Minimal generalization

Albright and Hayes (2003): There are *many rules* for the past tense, which we derive by combining more specific rules into more general ones

- Ø → əd / [vowt___]$_{[+past]}$
- Ø → əd / [nijd___]$_{[+past]}$
- Ø → əd / [X [+cor, +ant, −nas, −cont]___]$_{[+past]}$

We measure our confidence in a rule by how often it applies when its structural description is met ([glijd])

output	*rule*	*hits*	*failures*	*confidence*
glijdəd	Ø → əd / X{t,d}___	1146 (wɑnt)	88 (gɛt)	.929
glɛd	ij → ɛ / X{l,r}___d	6 (rijd)	1 (plijd)	.857
glowd	ij → ow / X C___C	6 (spijk)	178 (tijtʃ)	.033
glijd	– / X{t,d}___	29 (ʃɛd)	1205 (nijd)	.014

Islands of reliability are handled by specific rules with higher confidence than the general rule

Tolerance Principle

Yang (2016): children create a rule when, on average, it is more efficient to do so than to list everything

- If a rule applies to a word, it isn't listed as an exception
- Thus, before applying the rule, you have to check the whole list of exceptions (slower to make regulars than irregulars)
- The most efficient way to check the list is by frequency (speed of making irregulars correlates with frequency)
- If the list of exceptions is too long, the set cost of perusing the whole list of exceptions for frequent rule-abiders is too high to make up for the efficiency gained for less frequent rule-abiders

Under certain simplifying assumptions, this cost becomes too high when a rule that could apply to N words has at least $\frac{N}{\ln n}$ exceptions.

Tolerance Principle: vocabulary size

The smaller the number of items to which a rule could apply, the more exceptions are tolerated (Yang, 2016, p. 67):

items	*exceptions*	*% exceptions*
10	4	40%
100	23	23%
1000	145	15%

Thus, children's smaller vocabulary size may *help* them form productive rules

Tolerance Principle in the English past tense

Since many of the most frequent verbs are irregular, children must learn about 1000 verbs before creating the rule (Yang, 2016, p. 84):

number of most frequent verbs	*number of exceptions*	*number of allowable exceptions*
100	54	22
500	103	80
1022	127	147

Tolerance Principle in the English past tense

Since many of the most frequent verbs are irregular, children must learn about 1000 verbs before creating the rule (Yang, 2016, p. 84):

number of most frequent verbs	*number of exceptions*	*number of allowable exceptions*
100	54	22
500	103	80
1022	127	147

"Irregular" patterns are, at best, briefly productive: ɪ → æ / ___ŋ

verbs	*allowable exceptions*	*followers*	*exceptions*
3	2	ɹɪŋ, sɪŋ	bɹɪŋ
5	3	ɹɪŋ, sɪŋ, spɹɪŋ	bɹɪŋ, swɪŋ
8	3	ɹɪŋ, sɪŋ, spɹɪŋ	bɹɪŋ, swɪŋ, flɪŋ, stɪŋ, wɪŋ

This would explain why children occasionally say [bɹæŋ].

Tolerance Principle elsewhere

The artificial language experiment of Schuler et al. (2021) finds the cut predicted by the Tolerance Principle

- $\frac{9}{\ln 9} = 4.1$
- When one plural suffix applied to 5 of 9 nouns, children regularized it
- When one plural suffix applied to 3 of 9 nouns, children didn't

Artificial language experiments

Real languages aren't flexible enough to systematically answer all of our research questions

- Where is the line between Polish genitive singular and English past tense where children will (over)regularize?
- Do children and adults also impose order upon inconsistent outputs?

Artificial language experiments

Real languages aren't flexible enough to systematically answer all of our research questions

- Where is the line between Polish genitive singular and English past tense where children will (over)regularize?
- Do children and adults also impose order upon inconsistent outputs?

To answer questions like this, we can do experiments with *artificial languages*

- Make up words and sentences with the properties we want
- Teach adults and/or children our mini-language
- Ask them to make (given or new) words

When do children regularize?

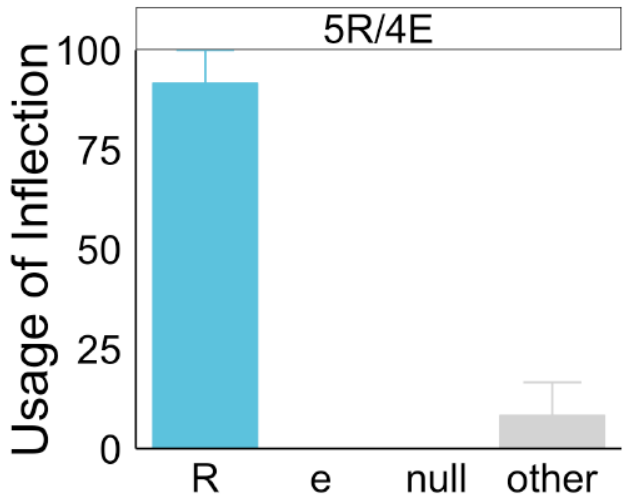
Schuler et al. (2016): American children aged 6–9 were taught nouns, then asked for the plural of new nouns:

singular	*plural*
mawg	ka
tomber	ka
glim	ka
zup	ka
spad	ka
daygin	po
flairb	lee
klidam	bae
lepal	tay
bleggin	???
norg	???
…	

When do children regularize?

Schuler et al. (2016): American children aged 6–9 were taught nouns, then asked for the plural of new nouns:

singular	*plural*
mawg	ka
tomber	ka
glim	ka
zup	ka
spad	ka
daygin	po
flairb	lee
klidam	bae
lepal	tay
bleggin	???
norg	???
…	

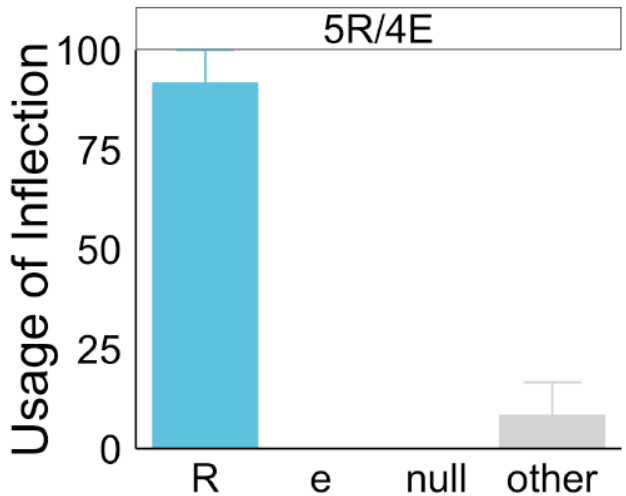


- They generalized the majority ending *ka* to all new words – regular rule

When do children regularize?

Schuler et al. (2016): American children aged 6–9 were taught nouns, then asked for the plural of new nouns:

singular	*plural*
mawg	ka
tomber	ka
glim	ka
zup	po
spad	lee
daygin	bae
flairb	tay
klidam	muy
lepal	woo
bleggin	???
norg	???
…	

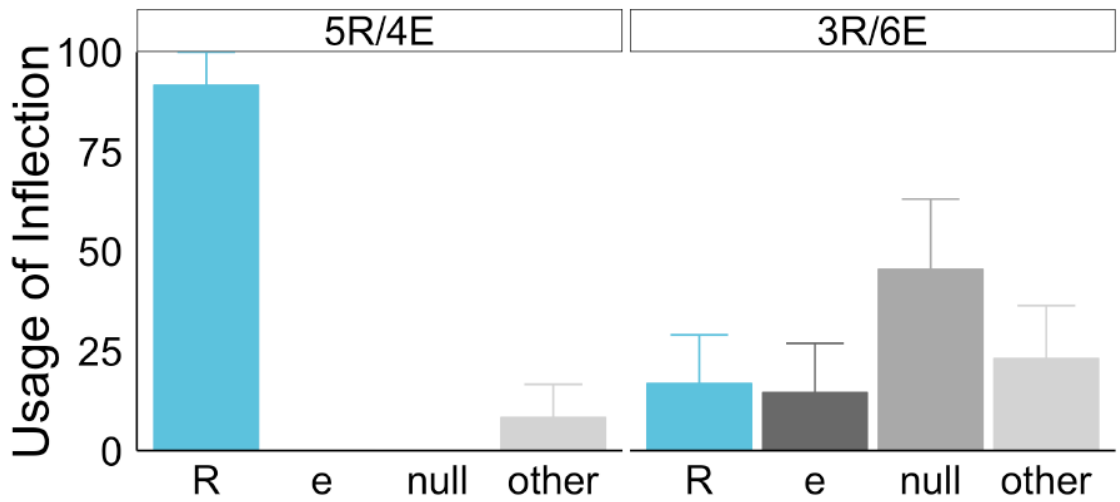


- They generalized the majority ending *ka* to all new words – regular rule

When do children regularize?

Schuler et al. (2016): American children aged 6–9 were taught nouns, then asked for the plural of new nouns:

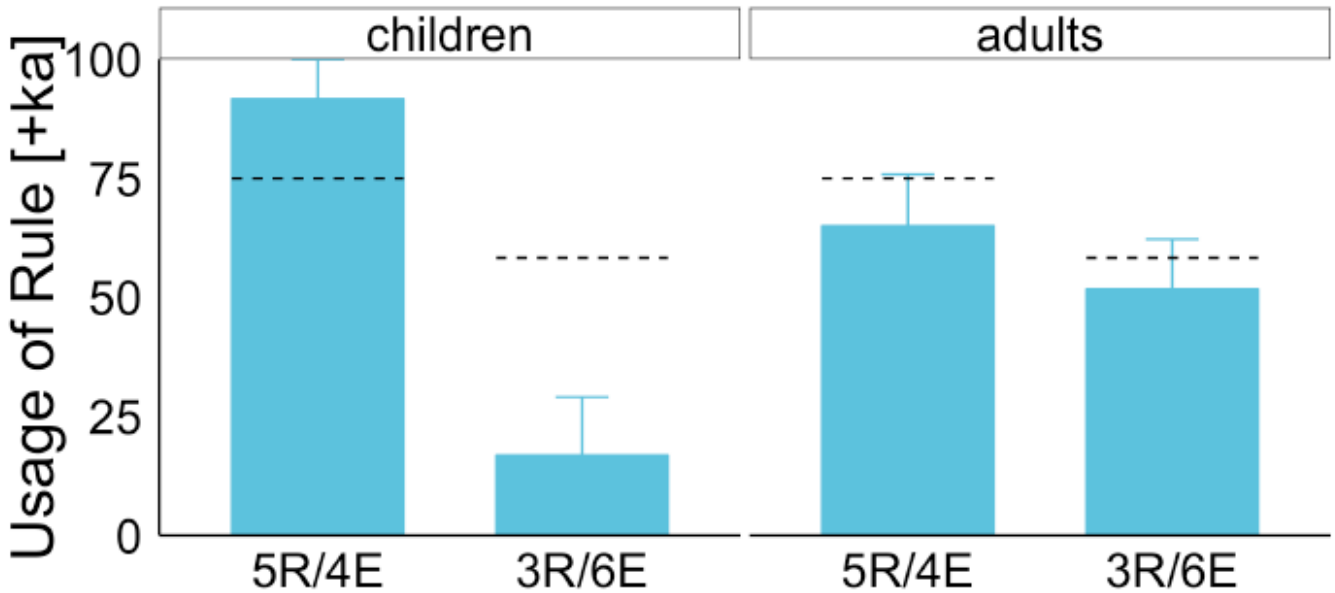
singular	*plural*
mawg	ka
tomber	ka
glim	ka
zup	po
spad	lee
daygin	bae
flairb	tay
klidam	muy
lepal	woo
bleggin	???
norg	???
…	



- They generalized the majority ending *ka* to all new words – regular rule
- When a minority of nouns had *ka*, chaos ensued – no regular rule

When do adults regularize?

When adults took the same study, they did something different:



Instead, they *matched the frequency* of *ka* from the sentences they heard

Outline

Frequency matching

In wug tests, adults tend to *match the frequencies* of their input – what can we learn from this?

- Speakers keep track of statistical patterns in their lexicon (mental dictionary of words/stems)
- Speakers may not pay attention to certain "unnatural" patterns

Attested frequency matching

What patterns can we find in wug tests?

- Assigning an underlying phoneme to a neutralized surface form (Ernestus & Baayen, 2003)
- Assigning a stem to a phonological class based on its phonology (Hayes & Londe, 2006)
- Assigning a suffix to a stem based on relevant phonological properties (Becker, 2009)
- Assigning a suffix to a stem based on irrelevant phonological properties (Gouskova et al., 2015)
- Assigning a suffix to a stem based on its other suffixes (Tabachnick, 2023)

References I

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*(2), 119–161.

Becker, M. (2009). *Phonological trends in the lexicon: The role of constraints* [Doctoral dissertation, University of Massachusetts Amherst].

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*(2–3), 150–177.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455.

Bybee, J. L., & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, *59*(2), 251–270.

Dąbrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. *Journal of Child Language*, *28*(3), 545–574.

Dąbrowska, E. (2008). The later development of an early-emerging system: The curious case of the Polish genitive. *Linguistics*, *46*(3), 629–650.

Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, *79*(1), 5–38.

Gouskova, M., Newlin-Łukowicz, L., & Kasyanenko, S. (2015). Selectional restrictions as phonotactics over sublexicons. *Lingua*, *167*, 41–81.

Hayes, B., & Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, *23*(1), 59–104.

Indefrey, P. (1999). Some problems with the lexical status of nondefault inflection. *Behavioral and Brain Sciences*, *22*(6), 1025.

References II

Kirov, C., & Cotterell, R. (2018). Recurrent neural networks in linguistic theory: Revisiting Pinker and Prince (1988) and the past tense debate. *Transactions of the Association for Computational Linguistics*, *6*, 651–665.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, *29*(3), 189–256.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (with Clahsen, H.). (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4).

McCurdy, K., Goldwater, S., & Lopez, A. (2020). Inflecting when there's no majority: Limitations of encoder-decoder neural networks as cognitive models for German plurals. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1745–1756). Association for Computational Linguistics.

Milin, P., Kuperman, V., Kostić, A., & Baayen, H. R. (2009). Words and paradigms bit by bit: An information-theoretic approach to the processing of inflection and derivation. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: Form and acquisition* (pp. 214–252). Oxford University Press.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, *28*(1–2), 73–193.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*(11), 456–463.

References III

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes, 8*(1), 1–56.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*: *Vol. 2. Psychological and biological models* (pp. 216–271). MIT Press.

Schuler, K. D., Yang, C., & Newport, E. L. (2016). Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. *Proceedings of the Annual Meeting of the Cognitive Science Society, 38*, 2321–2326.

Schuler, K. D., Yang, C., & Newport, E. L. (2021). *Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient* [doi:10.31234/osf.io/utgds].

Tabachnick, G. (2023). *Morphological dependencies* [Doctoral dissertation, New York University].

Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. MIT Press.

Zaretsky, E., & Lange, B. P. (2016). No matter how hard we try: Still no default plural marker in nonce nouns in Modern High German. In H. Christ, D. Klenovšak, L. Sönning, & V. Werner (Eds.), *A blend of MaLT: Selected contributions from the Methods and Linguistic Theories Symposium 2015* (pp. 153–178). University of Bamberg Press.