

The role of statistical learning in early L2 grammars: a generative perspective

Roger Hawkins
University of Essex

Objectives

- To report some observations of early L2 morphological development
- To describe a generative account of the observations with good empirical coverage, but some limitations
- To consider how learners' treatment of input (statistical learning) might be incorporated into an explanation

No knowledge of L2 → mental representations for:

- 1a. My name **is** X (3p copula, present)
- b. I **am reading** (1p auxiliary *be*, present/*-ing* progressive)
- c. She **eats** tofu (3p sing, habitual present)
- d. They **walked** over the hill (past)

Systematic optionality during development. L1 Chinese speaker describing a scene from *Modern Times*:

One day a girl **is walking** along the street. She **walking** past a baker shop and she **see** through the shop window. And er there **is** a lot of bread in the win, inside the shop. So she **feel** very hungry. ... And the girl **see** there **is** nobody around and she **stole** a loaf of bread and **run away** very quickly. But a woman **saw** her stealing a bread, so she **tell** the worker of the bread shop. She **tell** him 'er er a girl **stole** theirs bread'. Then the girl **ran away** an **bump** into a man.
Inflected forms of *be*/main verbs: *is walking, is (x2), stole (x2), saw, ran away*

Missing *be*/bare verbs: *walking, see (x2), feel, run away, tell (x2), bump*

Observations

I. Robust: forms of *be* supplied more frequently than affixal tense/agreement (Dulay & Burt 1973,1974; Pica 1985; Ionin & Wexler 2002; Haznedar 2007)

20 child L1 Russian → L2 English; 3yrs or less of immersion; age of first testing 3;9-13;10; 28 transcripts of conversations

Table 1. Frequency of suppliance of forms in obligatory contexts (based on tables 1 and 2 in Ionin & Wexler 2002)

	Cop <i>be</i>	Aux <i>be</i>	Reg <i>-ed</i>	3p <i>-s</i>
Freq	76%	63%	42%	21%
	329/431	300/479	73/174	67/321

I'. Less robust observation: asymmetry in suppliance of cop *be* vs aux *be*, past tense vs 3p *-s*

II. Non-target forms in English that relate to S-V agreement and simple past involve errors of 'omission' not 'commission':

She walking for *She is walking*, *The girl see* for *The girl sees*

not

**She am walking*, **They sees*

Table 2. Mis-use of forms (again based on Ionin & Wexler 2002, tables 1 and 2)

	Cop <i>be</i>	Aux <i>be</i>	Reg <i>-ed</i>	3p <i>-s</i>
Freq	8%	4%	0%	1%
	33/431	21/479	0/174	4/321

III. Early L2 learners produce a construction not encountered in input: *be* + bare v.

For NS *be* + v-ing expresses 'progressive' or 'future'; *be* + bare v has a wider range of meanings, e.g. (Ionin & Wexler 2002):

They **are help** people when people in trouble. (habitual)

He **is run away**. I stayed there. (past)

Table 3. Range of meanings of *be* + *bare V* (table 4 in Ionin and Wexler 2002: 112)

	Prog	Hab	Stat	Past	Fut	Ambig
Tokens	32	33	12	21	5	5
%	30	31	11	19	5	5

Even *be + v-ing* has a range of non-target meanings. Hawkins & Casillas (2007) examined one subject from the Ionin corpus: AY.

Table 4. Range of meanings of *be+v-ing*, *be + bare v* and *bare v-ing*: AY: Ionin corpus – CHILDES (MacWhinney 2000).

	Prog	Hab	Stat	Past	Fut	Ambig	Total
<i>be+v-ing</i>	11	6	0	6	2	3	28
<i>be+v</i>	1	10	3	7	2	3	26
<i>v-ing</i>	1	5	0	2	0	0	8

IV. Production of inflected forms increases gradually. E.g. Erdem – L1 Turkish; age 4 at start of recording; naturalistic exposure; 46 recordings over 18 months; inflected past tense verb forms in obligatory contexts (Haznedar 2007)

Table 5. No. of inflected past tense verb forms in obligatory contexts (based on tables 2 and 3 in Haznedar 2007:396)

Samples	8-12	18-31	32-39	40-46
Inflected v	15%	34%	40%	47%
	9/59	95/278	149/373	232/491

Summary of observations

I. Asymmetrical optionality in production:
Forms of *be* > *-ed*, *-s*.

In some studies: cop *be* > aux *be*
-ed > *-s*

II. In English, optionality involving S-V agreement and simple past = errors of omission, not commission.

III. *be* + *V*, *be* + *V-ing*, *V-ing* → non-TL meanings;
-ed and *-s* appear to be used with TL-meanings

IV. Optionality decreases incrementally – no sudden leaps.

A Missing Surface Inflection Hypothesis (Schwartz & Haznedar (1997), Prévost & White (2000) account of the observations

Verb morphology realises dependencies between the verb and other constituents with semantically interpretable features, e.g.

She [3p, +singular] Tense [-past] walk [3p, +sing, -past]
↓
walk s

Mechanism for capturing dependencies: uninterpretable features

v [*u*: person, *u*: number, *u*: tense]

u:Fs are valued by the interpretable features of constituents with which they merge

She walk s [v, *u*: 3p, *u*: +singular, *u*: -past]

She walk ed [v, *u*: 3p, *u*: +singular, *u*: +past]

I walk Ø [v, *u*: 1p, *u*: +singular, *u*: -past]

The valuing of *u*:Fs occurs in the syntax. The exponents of valued *u*:Fs are stored in a separate component (the Vocabulary), with contexts of insertion:

/s/	↔	/[v, -past, 3p, +singular]
/d/	↔	/[v, +past]
/Ø/	↔	/[v]

Where the contexts of insertion match the syntactic specification of *v*, the exponent is insertable

Notice that the Vocabulary entries are partly underspecified

/s/	↔	/[v, -past, 3p, +singular]
/d/	↔	/[v, +past]
/∅/	↔	/[v]

/s/ must be inserted when the verb's features are valued for 3p, +singular and –past. The verb takes the /∅/ inflection with any other subject where tense is -past

But underspecification means that /∅/ can be inserted in any context where /s/ (and /d/) can be inserted. How do we block:
**She walk along the street?*

An 'elsewhere' or 'subset' condition is required:

Forms in competition for insertion are subject to a 'subset principle' in native grammars (Halle, 1997): 'the item matching the greatest number of features specified in the terminal morpheme must be chosen.'

Claim of the MSIH:

- (a) L2 speakers have fully specified syntactic terminal nodes
- (b) Optionality arises because L2 learners do not obey the 'subset principle': default forms are inserted where a more specified form matches the terminal node

Example:

Vocabulary entries:

/s/ ↔ /[v, -past, 3p, -plural] + _____

/d/ ↔ /[v, +past] + _____

/∅/ ↔ /[v] + _____

Both v+∅ and v+s match the features of the terminal node:

[v, -past, 3p, -plural]

L2 learners sometimes select v+∅

Strengths of the proposal:

- (a) It provides an explicit mechanism for optionality in production
- (b) It predicts that when non-default forms are used they are used in a target-like way (e.g. no overgeneralisation of /s/, /d/)
- (c) L2 grammars are minimally different from those of native speakers = strong continuity in the availability of Universal Grammar

Limitations of the proposal

- Does not explain asymmetry of optionality between cop *be* and aux *be*, and between *-ed* and *-s* in those studies which find it.

/s/ ↔ /[v, -past, 3p, -plural] + _____

/d/ ↔ /[v, +past] + _____

/∅/ ↔ /[v] + _____

- Does not explain why *be* + bare v/v-ing/*be*+v-ing have a range of meanings not associated with the assumed input model (*be* +v-ing = 'progressive', 'future').
- Does not explain incremental increase in production of inflected forms over time.

Alternative proposal: a generative account that integrates certain input factors

There is empirical evidence that adult L2 learners remain good detectors of statistical tendencies in input.

- Saffran et al (1996): Adult learners can segment 'words' from a continuous stream of unknown syllables after 21 mins of exposure:

bi-da-ku-pa-do-ti-go-la-bu

- Unknown to participants: stream consisted of random combinations of 4 x 3-syllable 'words': *bidaku, padoti, tupiro, golabu*
- After 21 mins, participants showed a preference for 'words' over 'non-words': **da-ku-pa, *ku-pa-do*.

This suggests sensitivity to transitional probabilities between syllables, e.g. only *da* can follow *bi*; but *pa, tu, go* can follow *ku* = fast identification of discrete morphemes from experience.

Speculation: initially L2 learners do not have access to uninterpretable features

They can identify morphemes quickly from input, and assign them interpretable features.

This leads to the following hypothesis:

Missing Uninterpretable Features Hypothesis (MUFH)

Early L2 learners do not have access to *u*Fs. They assign 'contexts of insertion' to exponents on the basis of *i*Fs of terminal nodes with which they co-occur.

Early Vocabulary entries under the MUFH

NS: /s/ ↔ /[v, u: -past, u: 3p, u: -plural] + ____

Early L2 /s/ ↔ /[v] + ____ / [T, -past] ____ / [N, 3p, -plural] ____

NS: /d/ ↔ /[v, u: +past] + ____

Early L2 /d/ ↔ /[v] + ____ / [T, +past] ____

NS: /*(i)*s/ ↔ [T, -past, u: 3p, u: -plural]

Early L2 /*(i)*s/ ↔ [T] / [N, 3p, -plural] ____

Accounting for the acquisition of English verb morphology with the MUFH

- Asymmetric optionality: recall Ionin & Wexler's findings:

Frequency of suppliance

	Cop <i>be</i>	Aux <i>be</i>	Reg <i>-ed</i>	3p <i>-s</i>
Freq	76%	63%	42%	21%
	329/431	300/479	73/174	67/321

Claim: asymmetric optionality in production is a function of the *contextual complexity* of Vocabulary entries. Context-sensitive entries are computationally costly

Contextual Complexity Hypothesis (Hawkins & Casillas 2007)
The more nodes involved in stating a context of insertion, the more costly the entry and the less likely it is to be selected

/s/ ↔ /[v]+____ / [T, -past] ____ / [N, 3p, -plural] ____

/d/ ↔ /[v]+____ / [T, +past] ____

/(t)s/ ↔ [T]/[N, 3p, -plural] ____

see, walk ... ↔ [v]

- No misuse of inflections

Like the MSIH, the MUFH predicts that /s/ and /d/ will not be misused.

Unlike the MSIH, the MUFH offers a mechanism to account for asymmetric optionality **between** cop *be* and aux *be*, **between** simple past and S-V agreement:

/s/ is specified for more contexts than /d/; auxiliary *be* would *be* specified for more contexts than copula *be*

- Why do inflected forms increase in production over time?
Recall development of past tense marking in Erdem's speech (Haznedar 2007):

Samples	8-12	18-3	32-39	40-46
Inflected v	15%	34%	40%	47%
	9/59	95/278	149/373	232/491

Computational cost of context-sensitive Vocabulary entries is offset by frequency of activation of the entry.

Each activation of an entry increases the likelihood of its selection (another role for statistical learning in the grammar).

Selection of exponents will increase with exposure.

Summary of the proposal

- (a) Early L2 learners identify exponents of agreement quickly by computing 'transitional probabilities' in input.

- (b) MUFH = L2 learners assign 'contexts of insertion' to exponents of agreement on the basis of the \bar{F} s of co-occurring terminal nodes.

- (c) Asymmetric optionality follows from the computational cost of context-sensitive Vocabulary entries: the more nodes involved, the greater the cost of accessing an entry.

(d) Contexts of insertion for /s/, /d/ exclude misuse.

(e) Activation of entries over time offsets their computational cost (= incremental decrease in optionality).

(f) L2 grammars are minimally different from those of NS (initial absence of *u*Fs for analysing linguistic experience) = strong continuity in the availability of UG

Limitations

(a) The claim that u Fs are absent from early L2 representations requires independent justification.

(b) The claim that Vocabulary entries for exponents of agreement involve a specification of the \bar{r} Fs of co-occurring terminal nodes in the syntactic derivation requires independent justification.

Some evidence for (b): Early Vocabulary entries for verb morphology involve a specification of \bar{F} s of co-occurring terminal nodes in a syntactic derivation

Compare sentences like:

(a) My brother owns a house

(b) My best friend's brother owns a house

(c) The brother of my best friend owns a house

If selection of /s/ is determined by computing the \bar{F} s of co-occurring terminal nodes, /s/ should be equally easy/difficult to access in (a) and (b) (because the head of the subject NP is adjacent to the Tense category).

/s/ should be more difficult to access in (c) because a PP intervenes between the head of the subject NP and the Tense category.

Evidence from production

Hawkins and Casillas (2007) Speech production task

Participants see: stative or psych verb on computer screen:

own

Next they see a subject. One of:

The guest(s)

Or

The guest(s) of my music tutor(s)

Or

My music tutor's guest

Participants produce a complete sentence orally

Participants: 10 L1 Chinese, 10 L1 Spanish (all lower intermediate proficiency (Oxford QPT), 10 English controls

Results: **Table 6.** Mean % suppliance of V-s

	Expected	Chi	Spa	Eng
The guest	V-s	80	81	100
The guests	V-Ø	0	0	0
My music tutor's guest	V-s	79	82	100
The guest of my music tutor	V-s	60	61	100
The guest of my music tutors	V-s	47	49	100
The guests of my music tutor	V-Ø	40	35	0
The guests of my music tutors	V-Ø	1	0	0

Conclusions drawn from the findings:

- Complex subjects *per se* do not affect frequency of suppliance of –s

My music tutor's guest owns ... = The guest owns ...

- Participants not just selecting /s/ on the basis of the closest N. /s/ selected 20% less in

The guest of my music tutor owns ...

compared with

The guest owns .../My music tutor's guest owns ...

- Presence of a PP between the N subject and T/V reduces the selection of appropriate inflection on v
- At lower intermediate proficiency, L1 does not appear to play a role (Chinese = no S-v agreement, Spanish = S-v agreement)

Evidence from comprehension

Rationale

Low proficiency L2 learners should find it easier to identify subject number on the basis of N number marking than number agreement on v:

(a) /s/ ↔ /[N, +plural]+_____

(b) /s/ ↔ /[v]+____/[T, -past]____/[N, 3p, -plur]_____

(NB: This ignores the question of whether there is an independent Num category

Nagasawa 2007 (based on Johnson et al 2005)

Picture identification



Masked number: The duck-s-wims on the water
 The duck-s-swim on the water

Unmasked: The skunk eats the flowers
 The skunks eat the flowers

Participants: 10 intermediate prof. Japanese, 5 English controls

Results. **Table 7.** % accuracy in picture selection

Masked	Expected choice	Jap	Eng
v-s	Singular picture	32	100
v-∅	Plural picture	70	76
Unmasked			
v-s	Singular picture	90	92
v-∅	Plural picture	92	100

All subjects significantly different from chance except the Japanese speakers in the masked singular picture condition, where v-s is the crucial clue.

Conclusion: the Japanese speakers cannot compute the context-sensitive entry for 3p present singular /s/ in this task. In the unmasked condition they can compute the plural /s/ on the N.

Summary of the talk

- Adult L2 learners remain sensitive to input properties (transitional probabilities, semantically interpretable features of morphemes)
- Early L2 grammars represent exponents of agreement properties in terms of \bar{I} s of co-occurring terminal nodes (an effect of the *Missing Uninterpretable Features Hypothesis*)
- Context-sensitive Vocabulary entries are computationally costly (= optionality in selection)
- Statistical learning increases activation levels of context-sensitive Vocabulary entries (= incremental decrease in optionality)
 - More independent evidence required to support the claims

References

Dulay, H., Burt, M., 1973. Should we teach children syntax? *Language Learning* 23, 245-258.

Dulay, H., Burt, M., 1974. Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53.

Embick, D., Noyer, R., 2007. Distributed morphology and the syntax/morphology interface. In *The Oxford handbook of linguistic interfaces*, G. Ramchand and C. Reiss (eds), Oxford: Oxford University Press.

Halle, M., 1997. Distributed morphology: impoverishment and fission. *MIT Working Papers in Linguistics* 30, 425-449.

Hawkins, R., Casillas, G., 2007. Explaining frequency of verb morphology in early L2 speech. *Lingua* 1341.

Haznedar, B., 2007. The acquisition of tense-aspect in child second language English. *Second Language Research* 23, 383-417.

Ionin, T., Wexler, K., 2002. Why is 'is' easier than '-s'?: acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research* 18, 95-136.

- Johnson, V.-E., de Villiers, J.-G., Seymour, H.-N., 2005. Agreement without understanding? The case of third person singular /s/. *First Language* 25, 317-330.
- Lardiere, D., 2000. Mapping features and forms in second language acquisition. In *Second language acquisition and linguistic theory*, J. Archibald (ed), 102-129. Malden, MA: Blackwell.
- MacWhinney, B., 2000. *The CHILDES project: tools for analyzing talk*. 3rd edition. Mahwah, NJ: Lawrence Erlbaum.
- Nagasawa, M., 2007. Interpretation of 3rd person singular present tense –s by L2 speakers. MA dissertation, University of Essex.
- Pica, T., 1985. The selective impact of instruction on second language acquisition. *Applied Linguistics* 6, 214-222.
- Prévost, P., White, L., 2000. Missing surface inflection hypothesis or impairment in second language acquisition? evidence from tense and agreement. *Second Language Research* 16, 103-133.
- Saffran, J., Newport, E., Aslin, R., 1996. Word segmentation: the role of distributional cues. *Journal of Memory and Language*, 35: 606-621.
- White, L., 2003a. Fossilization in steady state L2 grammars: persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6: 129-141.
- White, L., 2003b. *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.