COGNITIVE APPROACHES TO SLL

1. INTRODUCTION

Universal Grammar-based researchers,
interested in the development of L2 grammars from a purely linguistic point of view:
emphasis on the *language* dimension of SLL
description/modelling of the linguistic systems → property theory
focus on competence
see language as a separate innate module

Cognitivists,
investigate hypotheses from the field of cognitive psychology and neurology
emphasis on the *learning* component of SLL
modelling the change/developmental processes of language acquisition → transition theory
how learners access linguistic knowledge; strategies they might employ; etc.
view SLA as one instantiation of learning among many others
believe we can understand the SLA process by understanding how the brain processes info

2. PROCESSING APPROACHES

scholars
concerned to develop transition theories
focus primarily on the computational dimension of language learning
may or may not believe language is a separate module

2.1. Information Processing Models of L2 Learning

how different memory stores deal with L2 information
how information is automatized and restructured through *repeated activation*
basic assumptions:

- Humans are viewed as active
- The mind is a general-purpose system
- Complex behavior is composed of simpler processes
- Component processes can be isolated
- Processes take time
- The mind is a limited-capacity processor

2.1.1. Attention-processing model

learners first resort to **controlled processing** in the L2 which
- involves the temporary activation of a selection of information nodes
- requires a lot of attentional control
- is constrained by the limitations of the short-term memory
- is typical of anyone learning a brand new skill

through repeated activation processing become **automatic**
automatized sequences

- stored in the long-term memory → available very rapidly, unconsciously, and effortlessly
- require minimal attentional control → automatic processes in parallel

Note: The distinction between controlled and automatic processing is one of routinization NOT one of conscious awareness

Attention to formal properties of language	*Information processing*	
	Controlled	*Automatic*
Focal	(Cell A) Performance based on formal rule learning	(Cell B) Performance in a test situation
Peripheral	(Cell C) Performance based on implicit learning or analogic learning	(Cell D) Performance in communication situations

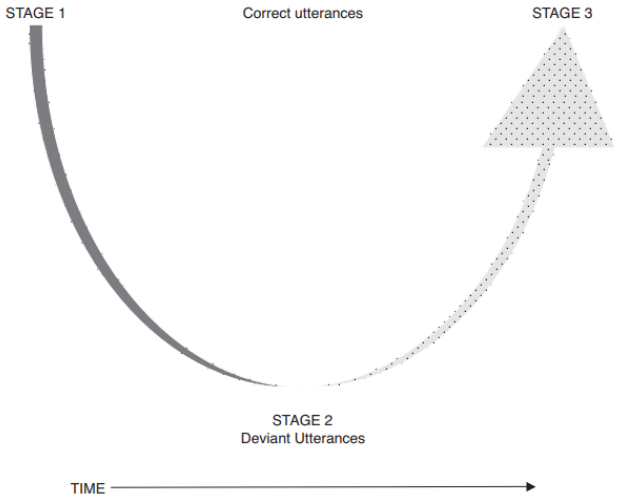
learning is movement from controlled to automatic processing via practice → controlled processes are freed to deal with higher levels of processing → incremental nature of learning
learning & **restructuring**:

learning is the inclusion of additional information → make changes to parts of the existing system the current system, i.e., restructuring

changes are discontinuous or qualitatively different from a previous stage

Time 1	Time 2	Time 3	Time 4
I am no go.	I am no go.	I am no go.	I am no go.
No look.	No look.	Don't look.	Don't go.
I am no run.	I am don't run.	I am don't run.	I am no run.
No run.	Don't run.	Don't run.	Don't run.

restructuring account for some of the variability of learner language: restructuring destabilizes some structures in the interlanguage:



2.1.2. Active Control of Thought (ACT) model

declarative knowledge (i.e., knowledge *that* something is the case) → **procedural** knowledge (i.e., knowledge *how* to do something)
essential differences between them:

- Declarative knowledge in an all-or-none manner VS. procedural knowledge can be partial
- Declarative knowledge acquired suddenly VS. procedural knowledge acquired gradually
- Declarative knowledge can be communicated verbally VS. procedural knowledge cannot

move from declarative to procedural knowledge:

- **Cognitive stage:**
  a description of the procedure is learnt, e.g. information from a teacher + concrete exemplars
  a lot of attentional control required
- **Associative stage:**
  a method for performing the skill is worked out
  proceduralization
    is achieved after a few trials
    reduces demands on working memory
  errors are likely during the associative stage
  knowledge at this stage is prone to restructuring
- **Autonomous stage:**
  the skill becomes more automatic
  extensive practice needed → decrease time required to perform the skill, error rate, etc.
  automatized knowledge outside attentional control → difficult to change or delete

2.1.3. Skill acquisition theory (SAT)

learning:
initial representation of knowledge → initial changes in behavior → eventual spontaneous, largely effortless behavior

- **Cognitive, declarative or presentation stage:** learners establish some new explicit knowledge
- **Associative, procedural or practice stage:** information about how to do something is put into action
- **Autonomous, automatic, or production stage:** a great deal of automatization has to take place through extensive practice

SLL → communicative practice serves as a device for proceduralizing knowledge of linguistic structures

practice needs to be skill-related → procedural knowledge is context-specific and cannot easily be transferred → uni-directionality or specificity of practice
declarative knowledge is transferable to other contexts → bi-directionality

problems of SAT:

- no explanation for the orders and sequences of acquisition
- insisting that all knowledge starts out in declarative form

six key SLA phenomena:

(1) *Why some structures never seem to enter the interlanguage at all?*
(2) *Why native-like forms are used in some contexts but not others?*
(3) *Why learning is incremental?*
(4) *Why there are differences between individual learners?*
(5) *Why there is fossilization?*
(6) *Why some structures are more likely to fossilize than others?*

2.2 Theories of Second Language Processing

Explore the factors controlling the way in which L2 learners process linguistic input

2.2.1 Input processing (IP)

working memory is capacity-limited → difficult to attend concurrently to different stimuli in the input ➔ main concern: how learners allocate attentional resources during online processing = what causes learners to detect certain stimuli in the input and not others

SLL: parse sentences → assign form-meaning relationships → comprehend utterances → convert L2 input into intake

linguistic data processed from the input and held in working memory for further processing

input processing does not offer

- a complete model of processing of input
- any explanation of how intake becomes integrated into IL system

a set of principles that explain the apparent failure of L2 learners to process linguistic forms:

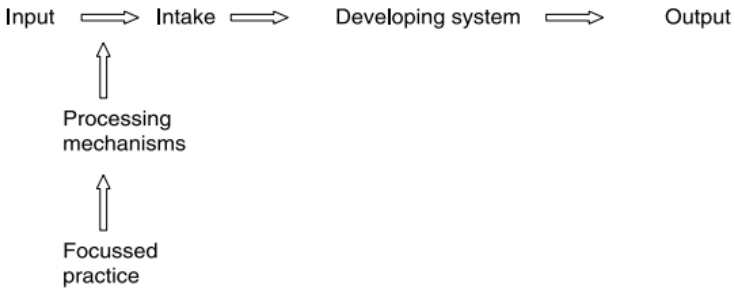
1) *The Primacy of Meaning Principle:*
   a) *The Primacy of Content Words Principle:* e.g., The **cat** is **sleep**ing.

b) *The Lexical Preference Principle:* e.g., I studi**ed** well **yesterday**.
c) *The Preference for Non-Redundancy Principle:* e.g., The cat is sleep**ing** vs. The **cat** sleep**s** ten hours **every day**.
d) *The Meaning Before Non-Meaning Principle:* e.g., Mary thinks **that** he is smart vs. He loves **that** girl.
e) *The Availability of Resources Principle*
f) *The Sentence Location Principle*

2) *The First Noun Principle*
a) *The Lexical Semantics Principle:* e.g., The fence **kicked** the horse
b) *The Event Probabilities Principle:* e.g., The child **scolded** the mother
c) *The Contextual Constraint Principle:* e.g., John is in the hospital because lo ataco Maria (**Mary attacked him**)

Processing instruction – explicit and input-based –: how learners process input and make form-meaning relationships → If we know what learners are doing wrong at the level of input processing, can we create pedagogical intervention to push them away from non-optimal processing?

Traditional approach to teaching grammar



IP approach to teaching grammar

→ IP attempts to deal with not just a linguistic difficulty, but with a problematic processing strategy

2.2.2. Autonomous induction theory (AIT)

Suzanne Carroll:

- A theory of the representation of language in the mind (i.e., property theory)
- A theory of how language is processed, both receptively and productively
- A theory of changes in linguistic competence (i.e., transition theory)

Jackendoff:

Integrative processors integrate primitives of some level of representation into complex structures → Carroll claims mental representations of language involve a number of distinct modules

Correspondence processors link the distinct autonomous representations → Carroll states their task is to map output onto a level higher up

a version of inductive learning (i-learning) → initiated when we fail to parse incoming language stimuli adequately using our existing mental representations = language acquisition device is triggered when the parsing system fails, e.g., *That's the cat whom the dog bit*

- **Note:** Adjustments lead to learning
- **Note:** Parsing presupposes linguistic competence; acquisition presupposes the absence of it

as children learn their L1, they develop processing procedures that are tuned to the specific grammatical properties of the L1 → language users reveal preferences for parsing sentences in particular ways

filter hypothesis

preferences transfer

fossilization

UG explains how a learner comes to have a representational system encoding phonological and morphosyntactic information → though children are not empiricist learners, adults learn L2 empirically

2.2.3 Processability theory

requirements:

processing component (based on Levelt's (1989) model of language generation)

theory of grammar (Lexical Functional Grammar, Halliday, 1985)

Pienemann believes:

language acquisition is the gradual acquisition of computational mechanisms → how learners acquire the computational mechanisms that operate on the linguistic knowledge

prod/comp of L2 forms can take place to the extent that they can be handled by the linguistic processor

learners must learn to exchange grammatical information across elements of a sentence → **feature unification**. Steps:

- identification of grammatical information in the lexical entry
- temporary storage of that information
- its utilization at another point in the constituent structure

the view on language production is based on Levelt (1989). Premises:

- processing components are generally not consciously controlled → normal, fluent speech
- processing is incremental
- output of the processor is linear, although it may not be mapped onto the underlying meaning in a linear way
- grammatical processing has access to a temporary memory store that holds grammatical info

the ability to match features across elements in a sentence develops gradually = learners have a **Hypothesis Space** which develops over time:

- level 1: lemma access, e.g., producing a simple word such as *play / Where is my book?*
- level 2: category procedure; lexical morphemes
- level 3: phrasal procedure, e.g., matching gender between *Det* and *N*
- level 4: simplified S-procedure; exchange of information from internal to salient constituent
- level 5: S-procedure; inter-phrasal morphemes; exchange of information between internal constituents, e.g., subject-verb agreement
- level 6: subordinate clause procedure

➔ learners will be able to share information across elements in a sentence in gradually less local domain

Teachability hypothesis

- L2ers follow a fairly rigid route in their acquisition of certain grammatical structures → *Stages of acquisition cannot be skipped through formal instruction*
- L2ers can only operate within their Hypothesis Space → *Instruction will be most beneficial if it focuses on structures from the next stage*

2.2.4. Nativization model

L2 acquisition consists of two general processes:

Nativization:

learners make the input conform to their own internalized view of what constitutes the L2 system ←→ they simplify the learning task by forming hypotheses based on knowledge that they already possess

Denativization:

learners accommodate to an external norm ←→ they adjust their interlanguage systems to make them fit with the input, making use of inferencing strategies

2.2.5. The efficiency-driven processor

provide explanations for language acquisition by reference to more basic non-linguistic factors
no innate linguistic constraints on the language processor

two cognitive systems:

Lexicon → as a repository of information about a language's words and morphemes, including information about:

category membership (N, V, etc.)
combinatorial propensities (co-dependencies)

co-dependencies could be thought of as the glue that holds a sentence together, e.g.

drink: V, <N N> → *Mary drank tea*.

how co-dependencies are to be resolved?

Computational system → operates on words, combining them in particular ways to construct phrases and sentences

these computational operations are carried out by working memory → the computational system should operate in the most efficient manner (i.e. *Minimize the burden on working memory*) → *co-dependencies must be resolved at the first opportunity*:

it strives to interpret incoming language and produce outgoing language so that as little information as possible needs to be stored. Production:

[Mary drank] → [Mary [drank tea]]

comprehension → processing costs are caused by:

having to revise an interpretation and so reactivate representations within WM
not knowing which elements of a sentence resolve a co-dependency
having an item left unresolved in the sentence

language acquisition is about the creation of mappings between form and meaning *Mary drank tea*

1. Interpret the first nominal (Mary)
   MARY
2. Access the meaning of the transitive verb *drink*; find its agent argument (MARY) to the left
   DRINK <agent: MARY>
3. Interpret the nominal to the right (TEA); treat it as the verb's patient argument
   DRINK <agent: MARY; patient: TEA>

Mary drank tea. ← form (sentence)
↓↑
DEANK <agent: MARY; patient: TEA> ← meaning

the same sequence of three operations is repeatedly activated → forming a **computational routine** which improves the speed and efficiency of the processor as they are gradually strengthened (**processing amelioration**)

development → as computational routines are formed
acquisition → as computational routines become entrenched
overgeneralization → as the computational routine of producing a structure (e.g. *–ed*) becomes automatized, making it less costly to overgeneralize than to inhibit the automatic routine
interference → as entrenched computational routines block or inhibit other routines developing (e.g. SVO to SOV)

3. EMERGENTISM

L2 learning is **bottom up** → language learning taps into the same, general, cognitive mechanisms that drive basic human learning in order to extract structures and patterns from the language input
usage-based view of development → knowledge of language is created and strengthened in response to opportunities to interpret and/or form utterances in the course of communication
learning → extraction of patterns from the language input they are exposed to → formal aspects of language **emerge** / are abstracted from language use, rather than being either innate, or learned as rules

Note: There is no way to explain how people come to know more than what they are exposed to

3.1. Input-Related Factors

what particular characteristics of L2 input can predict whether particular linguistic features are acquired early or late?
one characteristics is **cue** — animacy, case marking, agreement, etc. → **Competition Model**
functionalist/interactionist → learner's grammar results from the interaction between input and cognitive mechanisms

learner's task is to discover the particular form-function mappings
competition among various cues that signal functions like 'agent' (there is language-specific strength assigned to cues) – hence, competition model
speakers use four types of cues — word order, vocabulary, morphology, and intonation — to make form-function mappings:

e.g., *John kicks the ball* → cues are word order, animacy criterion, subject–verb agreement
e.g., *That teacher we like a lot* → cues are animacy criterion, case, subject–verb agreement

usefulness of a cue:

Cue availability → how often the cue is present when a particular pattern is being interpreted
Cue reliability → how often a cue points to a particular interpretation
Conflict validity → whether a cue wins or loses when it appears in competitive environments
In English, word order is a highly available & highly reliable cue for identifying subject
In English, agreement is highly reliable & often unavailable for identifying subject
→ there is language-specific strength assigned to cues

interference issue
the resolution of these conflicts

a) resort to L1 interpretation strategies
b) resort to a universal selection of meaning-based cues
c) gradually adopt the appropriate L2 biases as their L2 proficiency increases

learning → readjustment of which cues are relevant to interpretation and a determination of the relative strengths of those cues

3.2. Learner-Related Factors: Associative Learning CREED

Construction-based → learning involves learning and recycling *constructions*:
concrete lexical items such as *book*;
formulae such as *once upon a time*;
slot-and-frame constructions such as *give* [*someone*] [*something*];
more open abstract schemata such as [*subject V Obj1 Obj2*].

Rational → language representations in the mind are tuned to predict the linguistic constructions that are most likely to be relevant in the ongoing discourse

Exemplar-driven → language learning is formulaic (e.g., *Good+daytime*) ; acquisition is based on exemplars → every time the language learner encounters an exemplar of a construction, the language system compares it with memories of previous encounters

Emergent → regularities emerge as learners determine structure from language usage ; learning responds to and emerges out of the learner's experiences of the language rather than being the result of innately constrained rules

Dialectic → interaction with others or instructional events such as conscious learning help to mitigate the aspects of associative learning that cause L2 learning problems

3.2.1. Learners' use of frequency in the input

#1 the more times a stimulus is encountered, the faster and more accurately it is processed

#2 a critical component used in the calculation of probabilities:

humans have innate abilities to pick out cues and to calculate statistical probabilities implicitly → learning language is a statistical process: it requires the learner to acquire a set of associations between constructions and their function/semantic interpretations

frequency –i.e., positive evidence– does not guarantee noticing → for some features negative evidence may be needed

3.2.2. Overshadowing and attention blocking

Overshadowing

where a feature in the input is redundant → the feature may not be processed in a way that is useful for learning, e.g., *Yesterday, I carried that box.*

this processing failure is thought to happen when:

other features are more salient
world knowledge is likely to predict a communicatively satisfactory meaning
other features have been activated many more times before, in the L1

Attention blocking

overshadowing can over time lead to attention blocking:

if *x* has always expressed a particular meaning/function, it is difficult to associate *x* with another meaning or function, e.g. *-ly*

if a particular meaning has reliably been expressed using *x*, it is difficult to associate a different or an additional language feature with that same meaning, e.g. *indefiniteness*

3.2.3. Construction learning

language learning = learning constructions or conventionalized form-meaning mappings

are learners assisted by factors such as frequency or **prototypicality**?

verb locatives → *go somewhere*
verb object locatives → *put something somewhere*
ditransitives → *give somebody something*

3.2.4. Statistical learning and connectionist accounts

 # connectionist model = **parallel distributed processing (PDP)** (Rumelhart & McClelland, 1986) → neurally inspired models of human information processing: nodes connected by pathways

learning → as the network/learner is able to make associations between units:

Human mind is predisposed to look for associations between elements, (i.e., find regularities in the input = extract probabilistic patterns) and create links between them → knowledge emerges gradually, driven by the exemplars learners are exposed to

Associations become weakened through non-activation ≠ become strengthened through exposure to repeated patterns

implicit knowledge is conceptualized as a *complex adaptive system* that is in continual flux

the patterns of connections that emerge become so well-established that they reflect the categories and rules found in linguistic descriptions

what people learn and store in their implicit memories are 'memorized sequences' → acquisition occurs when statistical regularities are absorbed via implicit learning =

when learners unconsciously tally the likelihood that one form will follow another

capacity for language: ability to extract abstract categories from these memorized sequences by recognizing that a certain type of element occurs in a specific slot. This element then takes on an abstract value, e.g., *I don't …*

processes are unconscious: learners dwell on its products and formulate explicit rules → explicit knowledge can arise inductively

explicit knowledge can also be taught deductively

linguistic knowledge takes the form not of rules or items but of an elaborate system of weighted connections → rule-like behavior ≠ rule-governed behavior

the foregoing discussion on SLA theories could be subsumed under two paradigms:

Symbolists (information-processing theories & UG)

adopt an abstract view of linguistic representation → linguistic knowledge consists of a universal set of symbols and rules for combining them

make a distinction between property theory and transition theory

Connectionists

view linguistic knowledge as a complex network of associations → linguistic knowledge consists of associations of varying strengths, derived from elements encountered in the input

no distinction is make between representation (product) and learning (process) mechanisms = no belief in the distinction between competence and performance