Garrett’s Model

- Speakers engage in detailed planning before beginning to speak.
- Planning proceeds incrementally, in a cascaded fashion.

4 Stages
- Message Level
- Functional Level
- Positional Level
- Articulatory/Phonetic Level (Speech)

Support for Garrett’s Model

- Tip of the Tongue Phenomenon
  - Brown & McNeillage (1966)
  - “Would appear to be in a mild torment, something like the brink of a sneeze, and if he found the word his relief was considerable.”
  - TOT phenomenon indicates validity of distinction between functional & positional levels
- Speech Errors
  - Garrett’s theory predicts distinct & independent error types associated with different levels
- Word Errors occur at functional level
  - Should be sensitive to thematic and syntactic properties of words (aspects of the lemmas)
  - Should not be sensitive to information specified at the positional level, e.g. phonological form of lexemes

Experiments manipulate timing:
- picture and word can be presented simultaneously
- or one can slightly precede the other
- We draw inferences about time-course of processing

Schriefers, Meyer, and Levelt (1990)

- Auditory presentation of distractors
  - DOT phonologically related
  - CAT semantically related
  - SHIP unrelated word
- SOA (Stimulus onset asynchrony) manipulation
  - -150 ms (word …150 ms … picture)
  - 0 ms (i.e., synchronous presentation)
  - +150 ms (picture …150ms …word)

- Auditory presentation of distractors
  - DOT phonologically related
  - CAT semantically related
  - SHIP unrelated word

Schriefers, Meyer, and Levelt (1990)

Early Only Semantic effects
Late Only Phonological effects
Interpretation

- Early semantic inhibition
- Late phonological facilitation
- Fits with the assumption that semantic processing precedes phonological processing
- No overlap
  - suggests two discrete stages in production
  - an interactive account might find semantic and phonological effects at the same time

Speech production is at least in part an incremental process
Planning of complex sentences
- Meyer (1996): the arrow is next to the bag

- When you hear bow (sem.rel.1), uttering ‘the arrow is next to the bag’ is delayed
- When you hear suitcase (sem.rel.2), uttering ‘the arrow is next to the bag’ is delayed
- When you hear sparrow (phon.rel.1), uttering ‘the arrow is next to the bag’ is delayed
- When you hear bad (phon.rel.2), uttering ‘the arrow is next to the bag’ is not delayed

=> When starting to speak, not everything is ready, independent modules, can work at the same time!

Planning: X (fully) X+1 (partly)
Articulating: X

Constituent Structure in Generation

- Speakers generate language in phrases or constituents of phrases (clauses, NPs, VPs)
- Hesitations & Pauses
  - Mean pause length @ clause boundary = 1 s
  - Mean Pause length w/in clause = .75 s (Boomer)
  - Sentences planned one clause at a time
- Maclay & Osgood
  - Pauses at phrase boundaries filled by “Um,” “Ah”
  - Pauses within a phrase unfilled (Silence)
  - Utterances generated phrase by phrase

Constituent Structure in Generation

- When speakers repeat or correct themselves, they tend to repeat or correct a whole constituent
  - VP | NP | NP
  - Turn on the heater/ the heater switch.
  - Not: Turn on the heater/ on the heater switch.
  - VP | NP | NP
  - Turn on the stove/ the heater switch.
  - Not: Turn on the stove/on the heater switch.

Dell’s Model

- Dell
  - Semantic Level
  - Syntactic Level
  - Morphological Level
  - Phonological Level
- Garrett
  - Message Level
  - Functional Level
  - Positional Level
  - Articulatory/Phonetic Level

Dell’s Model

- Representations
  - exist at each of the four levels of model
  - processing typically more advanced at higher levels than lower levels
- Categorical Rules
  - set constraints on the categories and combinations of categories that are and are not acceptable
  - rules at each level define categories appropriate to that level
- Lexicon
  - network form w/nodes for words, morphemes, phonemes
- Insertion Rules
  - select the items for inclusion in the representation at each level
  - the most highly activated node belonging to the appropriate category is chosen
  - need verb, choose most active verb
  - once selected, item’s activation level immediately reduces to zero
Dell (1986)
- e.g., the semantic features mammal, barks, four-legs activate the word "dog"
- this activates the sounds /d/, /o/, /g/
- these send activation back to the word level, activating words containing these sounds (e.g., "log", "dot") to some extent

Explaining Speech Errors (Dell)
- Numerous nodes active at same time due to spreading activation
- Speech errors happen when activation spreads to the wrong item, such that it is more active than the target item

Dell’s explanation
- The process of making an error
  - The semantic features of dog activate “cat”
  - Some features (e.g., animate, mammalian) activate “rat” as well
  - “cat” then activates the sounds /k/, /ae/, /t/
  - /ae/ and /t/ activate “rat” by feedback
  - This confluence of activation leads to increased tendency for “rat” to be uttered
- Also explains the tendency for phonological errors to be real words
  - Sounds can only feed back to words (non-words not represented) so only words can feedback to sound level

Evidence for Dell’s model
- Mixed errors
  - Both semantic and phonological relationship to target word
    - semantic error = “dog”
    - phonological error = “hat”
  - Occur more often than predicted by modular models
    - if you can go wrong at either stage, it would only be by chance that an error would be mixed

Garrett & Dell on Error Data
- Spoonerisms
  Garrett reports 93% of spoonerisms within clause
  - Garrett – positional level
  - Dell – phonological level
- Word Exchange Errors
  I must let the house out of the cat.
  - Garrett – functional level
  - Dell – syntactic level
- Morpheme Exchange Errors
  He has already trunked two packs.
  - Garrett – positional level
  - Dell – morphological level
Dell vs. Garrett

- Closely Related
- Dell’s More Detailed
- Dell’s spreading activation neurally plausible and provides links to other cognitive processes

Predictions of Dell’s Model

- Errors belong to appropriate syntactic category
  - Also predicted by Garrett
  - Frequently true!
- Anticipation Errors
  Common
  The sky is in the sky.
- Anticipation errors turn into Exchange Errors
  I must write a wife to my letter.
- Anticipation errors involve short distances
- Lexical Bias Effect
  lewd rip rude lip
  2 x more common than:
  luke risk ruke lisk
- Speech errors can be multiply determined

Evidence for Dell’s Model

- Collections of Speech Errors
  - Mildly problematic…
- Speech Errors in the Laboratory
  - Different sorts of errors associated w/different deadlines (Semantic early/Phonological late)
  - More errors for rare words than frequent
  - Predicts speech errors for low frequency homonyms should be same as their high frequency counterparts

Conversational interaction

Conversation is more than just two side-by-side monologues.

Conversational interaction

Conversation is a specialized form of social interaction, with rules and organization.
**Conversation**

- Herb Clark (1996)
  - Joint action
    - People acting in coordination with one another
      - doing the tango
      - driving a car with a pedestrian crossing the street
    - The participants don’t always do similar things
  - Autonomous actions
    - Things that you do by yourself
  - Participatory actions
    - Individual acts only done as parts of joint actions

**Conversation**

- Herb Clark (1996)
  - Speaking and listening
    - Traditionally treated as autonomous actions
      - Contributing to the tradition of studying language comprehension and production separately
    - Clark proposed that they should be treated as participatory actions

**Conversation**

- Herb Clark (1996)
  - Speaking and listening
    - Component actions in production and comprehension come in pairs
  - **Speaking**
    - A vocalizes sounds for B
    - A formalizes utterances for B
    - A means something for B
  - **Listening**
    - B attends to A’s vocalizations
    - B identifies A’s utterances
    - B understands A’s meaning
  - The actions of one participant depend on the actions of the other

**Conversation**

- Herb Clark (1996)
  - Face-to-face conversation - the basic setting
    - Features
      - Immediacy
        - Co-presence
        - Visibility
        - Audibility
        - Instantaneity
      - Medium
        - Evanescent
        - Recordlessness
        - Simultaneity
      - Extemporaneity
      - Self-determination
      - Self-expression
    - Other settings may lack some of these features
      - e.g., telephone conversations take away co-presence and visibility, which may change language use

**Herbert Clark**

- Disfluencies aren't problems in speaking, but solutions to problems
- Disfluencies are signals that speakers plan in order to help coordinate their speaking with their listeners

**Language as Joint Activity**

- “Like waltzing, playing a duet, or shaking hands, it requires people to coordinate their individual actions in order to succeed (Clark, 1996).”
  - Ben considers projects Ann proposes Level 4
  - Ben understands what Ann means Level 3
  - Ben identifies Ann’s words Level 2
  - Ben attends to Ann’s voice as she vocalizes Level 1
Signals

Primary Signals
- Linguistic devices that enable Ann to communicate her message
  - Lexical semantics
  - Syntactic Structures
  - Suprasegmental Cues
  - Referential Gestures

Collateral Signals
- Lexical, syntactic, prosodic, and gestural devices that help coordinate primary signals
  - When she will vocalize
  - When she is about to revise or abandon an utterance

Pursue the Ideal Delivery
- Speakers try to produce utterances with ideal delivery
  - “the way they would have wanted to produce it if they had no problems (Clark & Clark, 1977).”
  - Characterized by standard prosodic theories
- Logic of Strategy
  - Listeners must attend to what speakers say
  - Processing is easier if expression arrives as expected
  - Speakers should produce utterances (or at least constituents) with predictable prosody

Signal the Initiation of Speaking
- In face-to-face conversation, speakers typically wait until they’ve established mutual gaze to begin
- Use of orienting expressions, e.g. “well”
  - Primary content: opposition
  - Collateral content: signals the initiation of speaking
- Produce pre-utterance filler, e.g. “uh” or “um”
- Produce first word and repeat it

Signal Your Intention to Suspend Speaking
- Since speakers rarely achieve the ideal delivery, they need signals to let listeners know when they will suspend speaking
- Nonreduced vowel
  - [ə] for a
  - [dhi] for the
    - [dhi] followed by suspension 81% LLC
    - Reduced ‘the” followed by suspension 6% LLC
  - [lə] for to
- Prolongation, [dhi:]

Signal Your Intention to Delay
- And, if possible, for how long.
- Delay signals
  - Uh signals a brief delay
  - Um signals a longer delay

Signal any Expression you Intend to Revise or Abandon
- Speakers have many techniques for signaling items to be revised
- Editing Expressions, e.g. “I mean”
  - Signals clarification/qualification
  - Content of resumption corresponds to the item that’s being clarified
  - Duncan: is there a doctrine about that, -- -- I mean a doctrine about u:h -- disfavoring American applicants,
- Prosody
  - Intonation of “they shortlisted” designed to match intonation of “they had”
  - Kate: they had. They shortlisted five people -- including me,
Do speakers really do this for the listeners’ benefit?

• "An alternative…is that they are not communicative acts, but simply the by-products of problems with planning utterances."

• Arguments against this:
  – Forms like uh and um are conventional
  – Planning does not require awareness
    • Selection of uh over um no different than the over a
  – Speakers can control their use of disfluencies
    • Good speakers don’t do these things in public speeches, but do do them in conversation

Meaning and understanding

• Common ground
  – Knowledge, beliefs and suppositions that the participants believe that they share
    • Members of cultural communities
    • Shared experiences
    • What has taken place already in the conversation
  – Common ground is necessary to coordinate speaker’s meaning with listener’s understanding

Structure of a conversation

• Conversations are purposive and unplanned
  – Typically you can’t plan exactly what you’re going to say because it depends on another participant
  – Conversations look planned only in retrospect
• Conversations have a fairly stable structure

Structure of a conversation

• Action sequences: smaller joint projects to fulfill a goal
  – Adjacency pairs
    • Opening the conversation
      – Kevin: Miss Pink’s office - hello
      – Joe: hello, is Miss Pink in
  – Exchanging information about Pink
    – Joe: is Miss Pink in
    – Kevin: well, she’s in, but she’s engaged at the moment…

Structure of a conversation

• Action sequences: smaller joint projects to fulfill a goal
  – Adjacency pairs
    • Exchanging the message from Worth
      – Joe: u’m Professor Worth said that, if Miss Pink runs into difficulties, … On Monday afternoon, … With the standing subcommittee, … Over the item on Miss Panoff, …
    • Kevin: Miss Panoff?
    • Joe: Yes, that Professor Worth would be with Mr Miles all afternoon, … So she only had to go round and collect him if she needed him, …
    • Kevin: ah, … thank you very much indeed,
    • Joe: right
    – Kevin: Panoff, right “you” are
    – Joe: right
    • Kevin: I’ll tell her,
    – Joe: thank you
    • Kevin: bye bye
    – Joe: bye
Opening conversations

- Need to pick who starts
  - Turn taking is typically not decided upon in advance
  - Potentially a lot of ways to open, but we typically restrict our openings to a few ways
    - Address another
    - Request information
    - Offer information
    - Use a stereotyped expression or topic

Opening conversations

- Has to resolve:
  - The entry time
  - Is now the time to converse?
  - The participants
  - Who is talking to whom?
  - Their roles
  - What is level of participation in the conversation?
  - The official business
    - What is the conversation about?

Identifying participants

- Conversation often takes place in situations that involve various types of participants and non-participants

Taking turns

- Typically conversations don’t involve two (or more) people talking at the same time
  - Individual styles of turn-taking vary widely
  - Length of a turn is a fairly stable characteristic within a given individual’s conversational interactions
  - Standard signals indicate a change in turn: a head nod, a glance, a questioning tone

Taking turns

- Typically conversations don’t involve two (or more) people talking at the same time
  - Three implicit rules (Sacks et al, 1974)
    - Rule 1: Current speakers selects next speaker
    - Rule 2: Self-selection: if rule 1 isn’t used, then next speaker can select themselves
    - Rule 3: current speaker may continue (or not)
    - These principles are ordered in terms of priority
      - The first is the most important, and the last is the least important
        - Just try violating them in an actual conversation (but debrief later!)
Negotiating topics

- Keep the discourse relevant to the topic (remember Grice’s maxims)
  - Coherence again
    - Earlier we looked at coherence within a speaker, now we consider it across multiple speakers
  - Must use statements to signal topic shifts

Closing conversations

- Closing statements
  - Must exit from the last topic, mutually agree to close the conversation, and coordinate the disengagement
    - Signal the end of conversation (or topic)
    - “Okay”
    - Justifying why conversation should end
      - “I gotta go”
    - Reference to potential future conversation
      - “later dude”

Summary

- “People use language for doing things with each other, and their use of language is itself a joint action.” Clark (1996, pg387)
  - Conversation is structured
    - But, that structure depends on more than one individual
    - Models of language use (production and comprehension) need to be developed within this perspective