Spoken English Production and Speech Reception Processes from Sentence Structure Perspective

Dr. Elsadig Ali Elsadig Elandeef [1] 👤 🆔 **and Dr. Ayman Hamad Elneil Hamdan** [2] 👤 🆔 ✉

[12]*Assistant Professor, English Department, King Khalid University, College of Sciences & Arts, Dhahran Aljanoub, Saudi Arabia*

✉ **Corresponding Author**: Dr. Ayman Hamad Elneil Hamdan, **E-mail**: aabdala@kku.edu.sa

ARTICLE INFORMATION	ABSTRACT
	This study aims to accentuate spoken production and speech reception regarding sentence formation. The study demonstrates the spoken production models such as Fromkin's Five Stage Model, The Bock and Levelt Model, Fromkin's Five Stage Model, Parallel –Processing Models and The Dell Model. It also states communicative problems strategies and many types of errors and mistakes relatively common in normal speech production, such as spoonerisms and speech errors. The study entails speech perception and how spoken language is perceived through linearity, segmentation, speaker normalization, and the basic unit of speech perception.

1. Introduction

Spoken production requires forming a conceptual representation that can be given in – linguistic form, then retrieving the right words related to that pre-linguistic message and putting them in the right configuration, and finally converting that bundle into a series of muscle movements that will result in the outward expression of the initial communicative intention. (Levelt, 1989). Furthermore, many autonomous components are responsible for different aspects of spoken production. These components include the conceptualizer, a component that is responsible for generating and monitoring messages; the formulator, in charge of giving grammatical and phonological shape to messages and which feeds on the lexicon; the articulator, which specializes in the motor execution of the message; an audition or acoustic-phonetic processor, which transforms the acoustic signal into phonetic representations; and the speech comprehension system, which permits the parsing or processing of both self-generated as well as other-generated messages (Meyer AS, 2000) . Spoken English language evolves from forming an idea in the speaker's mind before articulating it. The speaker constructs sentences from smaller parts or units that entail phones, phonemes, lexemes, phrase, clauses and sentences. English sentence production involves creating and expressing meaning through language. According to Levelt (1989), language production contains four successive stages: conceptualization, formulation, articulation, and self-monitoring. The conceptualization process requires deciding a targeted message that the speakers intend to convey. The message's decision without linguistic representation as an endpoint is known as preverbal message or message level of representation. Formulation involves the speaker must convert his or her message into linguistic forms. This stage involves lexicalization and syntactic planning. Lexicalization entails selecting the appropriate words, whereas syntactic planning enacts the words correctly and adds grammatical elements. Articulation or execution refers to the speaker must plan the motor movement needed to convey the message. [2]Once the speaker has organized his/ her thoughts into a linguistic plan, this information must be sent from the brain to the speech system's muscles to execute the required movements and produce the desired sounds from an articulatory phonetics perspective. Self-regulation is the last stage of speech production that refers to a

[2] (Scovel 1998:27). First, we must conceptualize what we wish to communicate; second, we formulate this thought into a linguistic plan; third, we execute the plan through the muscles in the speech system; finally, we monitor our speech, assessing whether it is what we intended to say and whether we said it the way we intended to.

set of flexibly used behaviours to guide, monitor and direct the success of one's performance. It is co-constructed within social interactions and influenced in various settings by others' attitudes and behaviors (Brown, 1983). Self-regulation includes three common sub-processes: self-observation or self-monitoring, self-judgment or self-evaluation, and self-reaction or behavioral adjustment (Griffin, 2003).

The production of spoken sentences involves the generation of a number of representation levels: a conceptual representation for the message wishing to convey, a grammatical representation that determines an appropriate word order for that message, and phonological and phonetic representations to guide articulation. Spoken sentences are based on putting words in a particular correct order embodying grammatical elements from both syntactic declarative and procedural knowledge and intuition. How the message is constructed is not conventionally acceptable among linguists, but common sense indicates that message is non-linear and must at least contain conceptual category information and has a thematic structure with concepts assigned to thematic roles (Allum, 2009). [3]Generation of sentences in spoken language production enacts that the speaker generates longer utterances, such as describing events or expressing emotions.

When speakers plan sentences, they retrieve words as described earlier. However, sentences are not simply a set of words but have a syntactic structure; speakers must apply syntactic knowledge to generate sentences. Core operation in speech production is preparing words from a semantic base. Sentence production entails a sequence of processing stages, beginning with the speaker's focusing on a target concept and ending with articulation initiation. The initial stages of preparation are concerned with lexical selection, zooming in on the appropriate lexical item in the mental lexicon. The following stages concern retrieving a word's morphemic phonological codes, syllabifying the word, and accessing the corresponding articulatory gestures.

2. Garrentt's Model of Syntactic Planning
According to Garrett's model, speech is produced linearly and that only one thing is processed at any one stage. At any one time in the course of a conversation, there would be more than one process taking place, such as when one is planning to what to say next while one is speaking. However, these different speech processes that occur concurrently are independent of one another and do not overlap. There are two major stages of syntactic processing, according to this model. One is at the functional level while the other is at the positional level. At the functional level, word order is not yet explicit. Words are semantically chosen and assigned syntactic roles such as subject and object. At the positional level, words are explicitly ordered.

Syntactic planning is dissociated from lexical retrieval because function and content words have different language production roles and are selected at different levels of the process. Content words are chosen at the functional level, whereas the selection of function words is made at the positional level. Garrett's theory predicts distinct and independent error types associated with different levels. Word Errors occur at a functional level; thus, speaker should be sensitive to thematic and syntactic properties of words (aspects of the lemmas), and he /she should not be sensitive to the information specified at the positional level, e.g., the phonological form of lexemes. Speakers generate language in phrases or constituents of phrases and their speech is interfered with by pauses at phrases boundaries filled by "Um," "Ah" Pauses within a phrase unfilled (Silence). When speakers repeat or correct themselves, they tend to repeat or correct a fundamental constituent. Many models are designed to study language production, such as:

2.1 Fromkin's Five Stage Model
Victoria [4]Fromkin was an American linguist who studied speech errors extensively. She proposed a model of speech production with stages that produced semantics, followed by syntax, and finally by phonological representation as follows: The intended meaning is generated; Syntactic structures are formulated; Intonation contour and placement of primary stress are determined;

---

[3] Following Garrett (1975), sentence production models generally assume that two distinct sets of processes are involved in generating syntactic structure (Bock & Levelt, 1994; Levelt, 1989). The first set, often called functional planning processes, assigns grammatical functions, such as subject, verb, or direct object, to lemmas. These processes rely primarily on information from the message level and the retrieved lemmas' syntactic properties. The second set of processes, often called positional encoding, uses the retrieved lemmas and the functions they have been assigned to generate syntactic structures that capture the dependencies among constituents and their order.

[4] Fromkin's (1971) Theory
-Utterance Generator model
-Top-down generator with 6 stages
– Generation of the meaning to be conveyed
– Mapping the meaning onto a syntactic structure
– Generation of the intonation contours of utterance
– Selection of words (content)
– Selection of words (function, affixes)

Word Selection-Content words inserted into the syntactic frame, function words and affixes added and phonemic representations added and Phonological rules applied (Meyer,2000).

2.2 The Bock and Levelt Model

This model[5] consists of four levels of processing. The first of which is the Message level, where the main idea to be conveyed is generated. The Functional Level is subdivided into two stages. The first, the Lexical Selection stage, is where the conceptual representation is turned into a lexical representation, as words are selected to express the desired message's intended meaning. The lexical representation is often termed the Lemma, which refers to the syntactical, but not phonological properties of the word. The Function Assignment stage is where each word's syntactic role is assigned. At the third level of the model, the Positional level, the order and inflexion of each morphological slot is determined. Finally, in the Phonological encoding level, sound units and intonation contours are assembled to form lexemes, the embodiment of the word's morphological and phonological properties are then sent to the articulatory or output system.

2.3 Parallel –Processing Models

In these non-modular models, information can flow in any direction and thus, the conceptualization level can receive feedback from the sentence and the articulatory level and vice versa. In these models, input to any level can therefore be convergent information from several different levels, and in this way, the levels of these models are considered to have interacting activity. Within a phrase, words that are retrieved initially constrain subsequent lexical selection.

2.4 The Dell Model

Dell's model of spreading lexical access activation is also commonly referred to as the Connectionist Model of speech production. Dell's model claims, unlike the serial models of speech production, that speech is produced by a number of connected nodes representing distinct units of speech (i.e., phonemes, morphemes, syllables, concepts, etc.) that interact with one another in any direction, from the concept level (Semantic level), to the word level (Lexical selection level) and finally to the sound level (Phonological level) of representation.

3. Speech Production Models

When one speaks, he /she needs to control a huge number of muscles, including the respiratory, laryngeal, and articulatory systems. In addition, many structures in these systems can move in different ways, at different speeds, and in different combinations. The speech motor system must somehow regulate all the speech subsystem's muscular contractions. [6]Speech production needs to consider the fact that sounds vary with the context in which they are produced and are influenced by speaking rate, stress, clarity of articulation, and other factors. Coarticulation is an integral aspect of speech production that results in enormous variability in producing a target sound. A given speech sound often can be produced in several different ways, and this variability in production is a central factor in speech motor regulation (Smith,2004).

3.1 Target Models

[5] The scope of lexical planning, which means how far ahead speakers plan lexically before they start producing an utterance, is an important issue for research into speech production, but remains highly controversial. The present research investigated this issue using the semantic blocking effect, which refers to the widely observed effects that participants take longer to say aloud the names of items in pictures when the pictures in a block of trials in an experiment depict items that belong to the same semantic category than different categories.As this effect is often interpreted as a reflection of difficulty in lexical selection, the current study took the semantic blocking effect and its associated pattern of event-related brain potentials (ERPs) as a proxy to test whether lexical planning during sentence production extends beyond the first noun when a subject noun-phrase includes two nouns, such as "The chair and the boat are both red" and "The chair above the boat is red". The results showed a semantic blocking effect both in onset latencies and in ERPs during the utterance of the first noun of these complex noun-phrases but not for the second noun. The indication, therefore, is that the lexical planning scope does not encompass this second noun-phrase. Indeed, the present findings are in line with accounts that propose radically incremental lexical planning, in which speakers plan ahead only one word at a time. This study also provides a highly novel example of using ERPs to examine the production of long utterances, and it is hoped the present demonstration of the effectiveness of this approach inspires further application of ERP techniques in this area of research.

[6] There is no model or set of models that can definitively characterize the production of speech as being entirely holistic (processing a whole phrase at time) or componential (processing components of a phrase separately). Despite their differences however, all models seem to have some common features. Firstly, the main question behind all models concerns how linguistic components are retrieved and assembled during continuous speech. Secondly, the models all agree that linguistic information is represented by distinctive units and on a hierarchy of levels and that the order in which these units are retrieved is sequential as they build upon one another. Thirdly, it seems that all models agree that you would need to access semantics and syntax prior to the phonology of an utterance, as the former dictate the latter and thus, all models share in common the following stages and substages in this order:
1) Conceptualization: deciding upon the message to be conveyed
2) Sentence formation:
a. Lexicalization: selecting the appropriate words to convey the message
b. Syntactic structuring: selecting the appropriate order and grammatical rules that govern the selected words
3) Articulation: executing the motor movements necessary to properly produce the sounds structure of the phrase and its constituent words

Target models describe speech production as a process in which a speaker attempts to attain a sequence of targets corresponding to the speech sounds he\she is attempting to produce (Indefrey,2011). Some theorists have suggested that these targets are spatial. Spatial target models posit that an internalized map of the vocal tract in the brain allows the speaker to move his or her articulators to specific regions within the vocal tract. The speaker can achieve the targets no matter what position the articulators begin the movement. The fact that articulators must reach a particular position from different starting points is important, because it means that the articulator's movements for a specific sound cannot be invariant but must change depending on the starting point.

3.2 Dynamic Systems Models
In this kind of theory, the degrees of freedom problem are addressed by positing that groups of muscles link together to perform a particular task. These linkages between muscles are not fixed: A muscle might be grouped with a particular set of muscles in what is called synergy or a coordinative structure to achieve one particular goal and with a different set of muscles in a different coordinative structure which refer to flexible groupings of muscles that may change depending on the particular speech output goal.

3.3 Connectionist Models
Computer models have been developed that simulate the human brain's neural processing. These models are also known as spreading activation models and parallel-distributed processing models (PDP). PDP models are based on a way of processing signals that is nonhierarchical. In other words, rather than finishing one step in the process before moving on to the next step, steps are processed more or less in parallel. This kind of processing is somewhat akin to how the brain processes information. Indeed, the performance of steps in parallel, or at least with much temporal overlap, is typical of speech production.

4. Sentence Production   and Message Formulation
Sentences are not born fully formed, but they are the product of a complex process. According to the standard view (Smith, 2004) sentence production spans over four independent sentence preparation stages: message, lemma, assembly, and articulation. Producing a sentence begins with creating a message – a conceptual representation of the event to be described linguistically. Then, the speaker translates the extracted message into an emerging sentence. This translation comprises stages of grammatical encoding of a sentence. Supposedly, grammatical encoding spans across two sub-stages: lemma retrieval, during which concepts receive their lexical names accompanied by their grammatical properties and grammatical assembly, at which the retrieved names assume positioning in the upcoming sentence. Finally, the speaker overtly produces the sentence at the stage of articulation. The production system in this and similar models is believed to be sequential and modular. It is sequential because processing at each preceding level has to be completed before processing at the next level can commence, and it is modular because processing at each level is believed to be encapsulated: for example, the speaker does not access lemmas at the message level or extract referential information at the assembly level. Access to the relevant information at each stage of sentence production is associated with accessibility statuses of the corresponding units. For example, at message level referents may receive a higher accessibility status due to their more conspicuous perceptual or conceptual properties (Hartley, 2001).   This may bias the speaker to process them earlier than the other referents when transferring the message details to the lemma level, affecting lexical accessibility of the words associated with these words' referent and grammatical properties. Suppose such preferential processing continues all the way to overt articulation. In that case, it is likely that the most accessible referent will be articulated before other referents taking part in the event and that it will be assigned as the most prominent grammatical constituent, for example, the Subject. This view helps understand how changes in accessibility at different production stages motivate the speaker's syntactic choices. In experimental settings, processing accessibility is often manipulated with the help of a priming paradigm (Griffin,2003).

The first component in Levelt's (as cited in Fromkin, 1998) production system is the conceptualizer. This component is responsible for generating the communicative intention and encoding it into coherent conceptual plans. In addition, the conceptualizer monitors what is about to be said as well as what has been said and how. In order to generate a message, declarative knowledge is accessed. Declarative knowledge includes encyclopedic knowledge (about the person's general experience of the world), knowledge about the situation (e.g. the interlocutor/s and the communicative context, among others), as well as information about the discourse record, that is, what has already been said. Levelt distinguishes two stages in message planning: macro planning and micro planning. It consists of retrieving information to express the sub-goals into which the overall communicative goal has been elaborated. In other words, it involves generating speech act intentions, like to narrate an event or express an opinion. The speaker's planning of a speech act, his selection of information to be expressed, and his linearization of that information are called macro-planning. Micro-planning divides that information into smaller conceptual chunks which are given the correct propositional shape and informational perspective. For instance, a small event's narration may be realized by a statement that can be presented in different ways.

In the next component in the production system, the formulator, the propositionally organized preverbal plan, activates the items in the lexicon that best correspond to the different chunks of the intended message that will, in turn, be responsible for transforming it into a linguistic structure. In Levelt's model and several other models, grammatical and phonological encoding are lexically driven. For grammatical encoding, both lexical access procedures and syntactic procedures are applied. In the lexicon, each lexical item is specified for semantic and syntactic information (lemmas), and morphological and phonological information (lexemes). Clark (1998) states that Human controlled processing tends to be serial, and is therefore slow. Conceptualizing a message requires a number of steps, such as constructing an internal representation, selecting the information to be communicated, breaking it into smaller chunks, and organizing them in a linear fashion shares processing resources with monitoring (Allum, 2009). Conversely, grammatical and phonological encoding are assumed to be automatic, which means that they do not require attention because they are single-step processes. According to Kempen (1987), the grammatical and phonological encoding of a message, including lexical articulation, are usually automatic and it can be concluded that parallel processing, incremental production, and automaticity allow for the speedy production of language in real-time.

5. Speech Perception
As with [7]speech production, many speech perception issues give direction to the theories attempting to explain how we analyze and perceive the spoken word. Some of these issues are linearity, segmentation, speaker normalization, and the basic speech perception unit.

5.1 Linearity and Segmentation
The linearity principle asserts that a specific sound in a word corresponds to a specific phoneme. The sounds that make up the word are distinct from each other and occur in a particular sequence. The segmentation principle is based perception. It is based on the notion that the speech signal can be divided into discrete units that correspond to specific phonemes (Bates, 2013). These two principles suggest that speech perception is based on a linear correspondence between the acoustic speech signal and the linguistic phonemic units. However, an abundance of research has established that this is not the case. In sum, speech perception theories emphasize different levels of processing of the speech signal. Decoding a spoken message involves the analysis of various-sized components of the signal, including acoustic, phonetic, phonological, lexical, suprasegmental, syntactic, and semantic components. Theories of speech perception can be categorized as active versus passive, bottom-up versus top-down, and autonomous versus interactive. Most theories of perception focus on acoustic-phonetic or phonemic aspects, including motor theory, acoustic invariance theory, direct realism, fuzzy logical models, and connectionist theories; recent theories also attempt to explain word recognition, including cohort theory and an interactive theory of speech perception (Zhao ,2013)

6. Communicative problems
Communicative problems that the speakers link to the variety of taxonomies of communication strategies existing in the problem-solving mechanism include the concept of communication strategy. Aristei (2011) proposes a framework that suggests a problem-solving mechanism for each type of problem. The main categories of problems have to do with:

i) resource deficits (e.g. because of an incomplete lexicon or insufficient morphological, or phonological specification);

ii) processing time pressure;

iii) perceived trouble in own output;

iv) perceived problems with the interlocutor's output.

Schnur and Costa (2006) suggest that speakers s can have difficulty retrieving lexical items from their incomplete L2 lexicon and grammatically and phonologically encoding their messages because the lexicon items are not sufficiently specified. When a lexical item cannot be retrieved three main options. Spoken language is conveyed via well-coordinated speech movements, which act as coherent control units referred to as gestures. These gestures and their underlying movements show several distinctive properties in terms of lawful relations among the parameters of duration, relative timing, range of motion, target accuracy, and speed. Unlike movements in locomotion or oculomotor function, speech movements, when combined into gestures are not mere physical instantiations of organs moving in space and time but, also, have intrinsic symbolic function.

---

[7] One view of speech perception is that acoustic signals are transformed into representations for pattern matching to determine linguistic structure. This process can be taken as a statistical pattern-matching problem, assuming relatively stable linguistic categories are characterized by neural representations related to auditory properties of speech that can be compared to speech input. This kind of pattern matching can be termed a passive process which implies rigidity of processing with few demands on cognitive processing. An alternative view is that speech recognition, even in early stages, is an active process in which speech analysis is attentively guided

Language-particular systems, or phonological grammars, are involved in these gestures' patterning. Grammar constraints regulate the permissible symbolic combinations as evidenced via eliciting judgments on whether any given sequence is well-formed in any particular language.

Furthermore, speech gestures are parts of words and thus one window into understanding the nature of the speech production1system is to observe speech movements as parts of words or larger chunks of speech such as phrases or sentences. The intention to produce a lexical item involves activating sequences of gestures that are part of the lexical item. The regulation in a time of the units in such sequences raises significant questions for speech motor control theories (but also for theories of cognition and sequential action in general). Major challenges are met in the inter-dependence among different time scales related to gestural planning, movement execution and coordination within and across domains of individual lexical items. How these different time scales interact and how their interaction affects the observed movement properties is, for the most part, still unknown.

7. Speech Errors and Mistakes
Many [8]types of errors and mistakes are relatively common in normal speech production. Errors and mistakes are categorized by specified mechanisms such as:

7.1Spoonerisms
[1]Errors in occur regularly and spontaneously in our speech. A case that is commonly known to most is that of Reverend Spooner. He has given his name to a particular type of error called spoonerisms that involves exchanging initial consonants between words. The following are examples of spoonerisms e.g.

 (You have **h**issed all my **m**ystery lectures.) Instead of (You have **m**issed all my **h**istory lectures.)

7.2Speech   Errors
Speech errors can be categorized according to the linguistic units involved in the error (i.e. at the phonological feature, phoneme, syllable, morpheme, word, phrase, or sentence levels) and the error mechanism involved (i.e. blend, substitution, addition, or deletion of units).

3.3 There are 8 main types of speech errors that occur at the different speech unit levels:
1. Phonemic segments
2. Phonetic features
3. Syllable
4. Stress
5. Morpheme
6. Word
7. Grammatical rules
8. Phrase

7.2.1 Phonemic Segments
There are four kinds of errors at this level: anticipation errors, preservation errors, and phoneme exchange and phoneme deletion.

7.2.1.1 Anticipation Errors
In anticipation errors, sounds which come later in the utterance inappropriately appear earlier than intended.
7.2.1.2 Consonant anticipation
a **r**eading **l**ist → a **l**eading list

---

[8] Types of errors and mistakes:Shifts: So, Homer phone Marge**s**
Exchanges: nife lite – hold card cash
 Anticipations: leading list – The currenth month is March
Perseveration: brushing his teesh
Additions: This class is getting slong...
Deletions: I'll get up during break and cape...
Substitutions: He'll never notice I'm present (missing)...
Blends: This prof is cruising for a breating...
Strands: I'm not in the read for mooding

hiring of the **f**aculty → **f**iring of the faculty
take my **b**ike → **b**ake my bike

7.2.1.3 Vowel anticipation
spl**i**cing from a t**a**pe → spl**a**cing from a tape
 7.3.1.4 Preservation Errors
In preservation errors, a sound produced early in an utterance reappears in an incorrect location later in the utterance.

7.3.1.5 Consonant preservation
a **ph**onological **r**ule → a phonological **f**ool
he **p**ulled a **t**antrum → he pulled a **p**antrum

7.3.1.6 Vowel preservation
annotated b**i**bliography → annotated b**a**bliography

7.3.1.7 Phoneme Exchange
In phoneme exchange errors, phonemes are reversed.
b**r**ake f**l**uid → b**l**ake f**r**uid
f**i**ll the p**oo**l → f**oo**l the p**i**ll
B**e**db**u**gs → b**u**db**e**gs

7.3.1.8 Phoneme Deletion
In phoneme deletion errors, phonemes are deleted.
Speech error → peach error

7.3.2 Phonetic Features
Phonetic features are the attributes that combine to define the phonemes of a language.

7.3.2.1 Voicing reversal
(a) **C**lear **b**lue sky → **g**lear **p**lue sky (voicing reversal; voiceless → voiced)
(b) **P**at a **d**og → **b**at a **t**og (voicing reversal; voiceless → voiced)
In example (a), the /k/ voiceless velar feature in the word *clear* is exchanged with the voiced feature of the /b/ voiced bilabial stop and becomes /g/ voiced velar and /p/ voiceless bilabial stop accordingly.
In example (b), the /p/ voiceless bilabial stop feature is exchanged with the /d/ voiced alveolar feature and becomes /b/ voiced bilabial stop and /t/ voiceless alveolar feature.

7.3.2.2 Voicelessness anticipation
He's a **v**ile **p**erson → He's a **f**ile **p**erson
In the example above, /v/ is voiced labiodental fricative and /f/ is voiceless labiodental fricative. The voicing feature changes in anticipation of /p/ voiceless bilabial stop that follows.
 7.3.2.3Nasality reversal
Cedars of Leba**n**on → cedars of Le**m**adon

7.3.3Syllable
In syllable errors, there is a structural flaw that is observed with regard to the syllable place. Initial segments **replace** initial segments, and final segments replace final segments.

7.3.3.1Syllable-deletion
Una**ni**mity of opinion → una_mity of opinion
Mutter **un**intelligibly → mutter _intelligibly

7.3.3.2 Syllable exchange
Stockwell and **Schact**er → **Schac**well and **Stock**ter

7.3.4 Stress
Apples of the **O**rigin → Apples of the o**RI**gin

mo**BIL**ity → mobi**LI**ty

7.3.5 Word Selection and Placement Errors
7.3.5.1 Word Exchange
tend to turn out → turn to tend out

7.3.5 .2 Word Movement
I **really must** go → I **must really** go

7.3.6 Morpheme
7.3.6.1 Inflectional morpheme errors
Rules of word formation → words of rule formation
I'd forgott**en** about that → I'd forgot about**en** that

7.3.6.2 Derivational morpheme errors
easi**ly** enough → easy enough**ly**
There's a good likeli**hood** → There's a good likeli**ness**

7.3.7 Grammatical Rules
The last I **knew** about it → The last I **knowed** about it
The **children** are in the park → The **childs** are in the park

7.3.8 Phrase
A hummingbird was attracted by **the red colour** of the feeder
↓
The red colour was attracted by **a hummingbird** of the feeder

8. Conclusion
Many factors may implicate speech productions that are not incorporated into most models, such as emotion and physical condition. No model or set of models can propose that the production of speech is entirely holistic or componential. Despite their differences, all models seem to have some common features: models concern how linguistic components are retrieved and assembled during continuous speech. The models agree that linguistic information is represented by distinctive units and on a hierarchy of levels, and it seems that all models agree that one would need to access semantics and syntax prior to the phonology of an utterance, as the former dictates the latter. Language production contains four successive stages: conceptualization, formulation, articulation, and self-monitoring. All models share in common the following stages: conceptualization: deciding upon the message to be conveyed; Sentence formation; Lexicalization: selecting the appropriate words to convey the message; Syntactic structuring: selecting the appropriate order and grammatical rules that govern the selected words. Articulation: executing the motor movements necessary to properly produce the phrase's sounds structure and its constituent words.

References
[1]   Allum, P. H., & Wheeldon, L. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual–syntactic interface. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(5), 1240.
[2]   Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: the role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 791.
[3]   Aristei, S., Melinger, A., & Abdel Rahman, R. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of cognitive neuroscience*, *23*(7), 1567-1586.
[4]   Bates D., Maechler M., Bolker B., Walker S. (2013). lme4: Linear Mixed-Effects Models Using Eigen and S4. R Package Version 1.04. Available at: http://CRAN.R-project.org/package=lme4
[5]   Brown, G. & G.Yule. (1983). *Teaching the Spoken Language*. CUP
[6]   Levelt, W.J.M. (1989). *Speaking: from Intention to Articulation*. Cambridge,Mass: MIT.
[7]   Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive psychology*, *37*(3), 201-242.
[8]   Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57-84.
[9]   Fromkin, V.A. & Ratner, N.B. (1998). *Speech Production*. In J.B. Gleason & N.B. Ratner, (Eds.), Psycholinguistics (2nd ed.)(pp.309 – 346). Belmont: Thomson Wadsworth.
[10]   Griffin, Z. M., & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language*, *49*(4), 537-555.
[11]   Hartley, T. (2001). *The Psychology of Language* (2nd ed.). East Sussex: Psychology Press.

[12] Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in psychology*, *2*, 255.

[13] Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive science*, *11*(2), 201-258.

[14] Meyer, A. S., & Van der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin & Review*, *7*(2), 314-319.

[15] Schnur, T. T., Costa, A., & Caramazza, A. (2006). Planning at the phonological level during sentence production. *Journal of psycholinguistic research*, *35*(2), 189-213.

[16] Smith, M., & Wheeldon, L. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 675.

[17] Wheeldon, L., Ohlson, N., Ashby, A., & Gator, S. (2013). Lexical availability and grammatical encoding scope during spoken sentence production. *Quarterly Journal of Experimental Psychology*, *66*(8), 1653-1673.

[18] Wheeldon, L. (2013). Producing spoken sentences: The scope of incremental planning. *Speech production and perception: Speech planning and dynamics*, 97-118.

[19] Zhao, L., & Yang, Y. (2013). Grammatical planning scope in sentence production: Evidence from Chinese sentences. *Acta Psychologica Sinica*, *45*(6), 599-613.