

Off-line measures in Human Speech Production

Seminar in Psycholinguistics SS 2004

Renjini Narendranath

Dept. of Computational Linguistics
Universität des Saarlandes
renjinin@coli.uni-sb.de

October 27, 2004

Abstract

In this term paper we present the results of three important papers on speech production. Though the papers agree on the nature of the production proceedings, they have serious disagreements on the fundamental natures of the kind of interactions between these modules. Specifically, these papers argue for and against the independence of the stages of the grammar as well as the phonological production modules. The experiments to prove were taken from the everyday speech errors like slip-of-the-tongues where people mispronounces some words resulting in the formation of nonsensical or legitimate words and errors like tip-of-the-tongue states where the speaker reports that he/she knows the word but cannot retrieve the phonological form of the word to articulate.

1 Introduction

In this report, we analyze and discuss the various off-line measures in Speech Production. We also have a brief sketch of the experiments and their results with regard to various parameters of the off-line measures of speech production. Off-line measures are those that takes place inside the brain from the conception of the message to be communicated till it results in overt speech. The errors in any of this lets say “preparation” stage before articulation, would result in speech errors like slip-of-the-tongues or tip-of-the-tongues. The study of such errors during speech provides informations about the way humans think and speak (refer 2 for Chomsky’s Competence and Performance). Psycholinguistics as a field of research offers explanations about the way human beings think and speak. This field of study among others uses different sources for study. It not only analyses the errors in speech, but also errors in writing.

We discuss here three papers that deals with the investigation of how different speech errors point to the different levels/stages of speech production. The important question that gets addressed in these papers are whether the off-line processes that sometimes cause errors, are dependent or independent of each stages of speech production as outlined in the section : 3.

Organization of this paper

In the section : 2, we explain some terminologies with respect to the themes discussed in this report. We further discuss an off-line model of speech in section : 3. In the next section : 4 we concentrate on the interactive model of speech production as outlined in the Dell and Reich paper [DR81]. We briefly discuss the views of other researchers on the lexical bias in sound errors. We also discuss the different experiments carried out with the Toronto corpus of speech errors. In the coming section : 5 we see how the gender ascription tendencies during tip-of-the-tongue stages among Italian speakers, account for the independence of the two stages of speech production. In the section : 6 we describe briefly the experiments of Miozzo and Caramazza on the lexical-syntactic retrieval in tip-of-the-tongue states. We make a comparison and discuss the results of these three papers in the section : 7. And finally, we conclude this report in the section : 8.

2 Some Basic Terminologies

- Offline measures: During speech production, there are many thought processes that happens in the background as a result of interactions among the different modules.
- Linguistic competence and Performance : In Chomskian tradition, the study of linguistics can be classified into two broad classes of description. Linguistic competence that deals with the competence of a person to produce syntactically correct sentences, and linguistic performance that deals with the thought process behind conversations (written as well as spoken).

3 Off-line model of Speech Production

That there exists a distinction between semantic, syntactic and phonological information about the spoken words, is a widely accepted fact by the speech production researchers. They also have agreed on the fact that not only these informations are stored differently, they are also accessed independently.

The first one that we are going to look at is was developed by Garret. According to him, in speech production, Figure: 1 illustrates this idea of linear process.

In 1975, Garret proposed that the sentence production process contains two independent and serially ordered stages. One is the *functional stage* and the other is *positional stage*. *Functional stage* is where the underlying grammatical representation of the sentence is constructed and *positional stage* is the place where the phonological informations of the words to be spoken is represented. Figure:1 represents this model of serially ordered levels of sentence production.

This proposal was made on the basis of studies about the slips of the tongue errors. Table: 4 illustrates some of such errors during speech production. Garret argues that these errors have something in common. That is, in all these errors, there are some elements that interact with each other and cause errors. In the case of sound exchanges, the interacting elements (IE) are the reversing units. For substitutions, anticipations and preservations, the IE are the target element and the element displacing it in the error. In the case of blends, the two “blending words” are the IE.

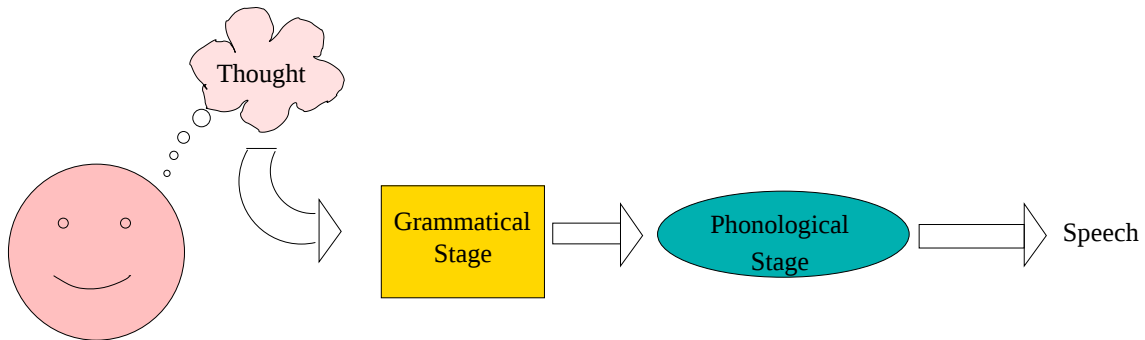


Figure 1: A linear-modular view of speech production

Garret argues that, the functional stage and the positional stage are independent because of the following three reasons:

1. The grammatical class of the words comprising or containing the interacting elements.
2. The phonological similarity of the elements and their environments.
3. The span or distance over which the elements move.

The errors 1–5 in the table : 4 are supposed to have occurred in the positional stage. The IE's in these errors are phonologically similar and often belong to different grammatical categories, also often their positions are adjacent to each other. The errors 6–8 are of different nature. The IE's here are not phonological similar, but they do belong to the same grammatical class. In some cases, the IE's are far apart in contrast to the other positional stage errors. Therefore, these errors might have occurred in the functional stage, i.e, when they were encoded grammatically. Garret thus argues that, the errors happening in each stages shows difference in their very nature: the errors occurring in the positional stage are constrained by the proximity and phonological similarity. But the functional errors are influenced only by the grammatical similarity. That is, the two stages: positional and functional stages are blind to each other, and do not interact with each other.

4 Stages in sentence production from the analysis of speech error data

In this section we will discuss about the hypothesis formulated in Dell and Reich's paper [DR81]: during speech production, whether there is any interaction between the positional and functional stages. We mentioned in the section:3 that it has already been accepted that there exists three different linguistic modules through which informations flows till it is actually produced as an overt speech. Dell and Reich have argued in favor of this structure but they questions the argumentation made by Garret that the positional and functional stages are blind to each and that they do not interact with each other during sentence production. To be specific, Dell and Reich tested the following two hypothesis:

- Whether or not the positional stage is blind to the sort of information that is normally accessed only at the functional stage. This was tested by determining whether or not sound misordering errors tend to create legitimate words.
- Whether the functional stage is blind to the informations accessible to the positional stage (phonological representations). This was tested by analyzing the phonological similarity of the IE's in functional errors.

Error / Slip of the tongue		
1.	Sound Exchange :	York library <i>lork yibrary</i>
2.	Sound Preservation :	beef noodle <i>beef needle</i>
3.	Sound Anticipation :	reading list <i>leading list</i>
4.	Bound morpheme shift :	they get weirder everyday ... <i>weird everier</i> ..
5.	Combined-form exchange :	talking Turkish <i>Turking talkish</i>
6.	Word misordering :	Writing a letter to my mother Writing a <i>mother</i> to my <i>letter</i>
7.	Word blends :	Hail a taxi/cab Hail a <i>tab</i>
8.	Word substitution:	On my knee On my <i>elbow</i>

The evidence to support these hypothesis was taken from the analysis of everyday speech errors. The table : 4 illustrates some such slips of the tongues that we tend to commit in everyday speech.

Hypothesis 1 : Positional stage blind to functional stage ?

As opposed to Garrets argumentation that positional level does not have access to the lexicon (where the grammatical aspects of the words are stored), Dell and Reich explains that, the errors like:

pitch fork → *fitch pork*

occurs by creating legitimate word strings. They call this tendency of errors to create words Lexical Bias. The error of the sort mentioned above according to Garret belongs to the positional stage error. Dell and Reich argues against this by stating that, if the error occurs at the positional stage (which according to Garret is blind to the functional stage informations of the lexicon) then how does such an error produces a legitimate word. That is, the positional stage at some point of time must have accessed the mental lexicon and got the word, for example: *pork* . This lexical bias, for creation of new

words, clearly favors the argumentation that the positional stage does access the lexicon which is also accessed by the functional stage. This argument challenges Garrets results that positional stage and functional stages are blind to each other.

Here are some examples of Spoonerism:

“ Modeling dialogues” → ‘*Dialing modelogues*’
 “You missed my history classes.” → “You *hissed* my *mystery* classes.”
 “We seem to have held our head fairly evenly with other colleges.” →
 “We seem to have held our *heaps* evenly with other colleges.” (taken from [Pot80])

4.1 Views of other researchers on Lexical Bias

Victoria. A. Fromkin (1973) and D. B. Fry (1977) found that most of the phonemic slips create nonsensical words. Therefore, the positional stage is independent from other stages. We are not going into detail of the kind of corpus or investigation materials they have used to make this claim.

Sigmund Freud (1924) in contradiction to Garret, Fromkin and Fry considered that, errors will not only tend to create new words, but also that such words would reflect the speaker’s thoughts and inner motives.

A. A. Hill (1972) who supported Freud’s argument remarks that “Spoonerism” nearly always created at least one word form in his experience. Another supporting argumentation for lexical bias in sound errors, came from Baars et.al’s work. In their experiments, they artificially created potentially spoonerizable words and them with the error rates. Almost 75% of slipped words were valid words showing a Lexical bias and the rest 25% created non-words. Kempen, however noted that the majority of word outcomes could reflect a post encoding editorial check that vetoes non-word outcomes, and that this editorial process could only be a product of the induction procedures and not a part of normal production. That means the 75% lexical bias found in Baars et.al’s experiment could be actually resulting from the editorial feedback between the processes.

Garret (1976) made some preliminary inquiries about the evidence for genuine lexical bias in naturally occurring errors. On the basis of the corpus used by him as well as Fromkin, Garret concluded that such lexical biases are less than 41%. This leads to a serious disagreement problem and thereby creating a need for analyzing the issue of lexical bias again. In Dell and Reich’s experiments, they analyzed the tendency for sound errors in a new corpus to make words.

4.2 Results of the Toronto Corpus based Experiments

In the experiments, Dell and Reich present the tendency for sound errors in a new corpus to make words, and compare the proportion of word outcomes with an estimate of how often word outcomes should result by “chance”. The *Toronto Corpus* which they used for this analysis contained naturally occurring errors from approximately 4000 English Examples collected in Toronto by about 200 students of psycholinguistics. This corpus was made in a period of two years (1975–77) and are restricted to the word level errors and sound errors involving only initial consonant changes.

Experiment on small corpus

In the following, we review the estimation of chance as outlined in the Dell and Reich paper. A sample corpus of four exchange errors was used for this explaining the probabilities. In reality, the procedure was done once with 196 exchange errors and once with 99 participants to obtain the base line probabilities.

A sample corpus of four exchanges:

pitch fork → fitch pork
Lawrence the Warden → Wawrence the Larden
postal code → coastal poded
chin tickled → tin chickled

STEP 1: Create two lists

<i>List 1</i>	<i>List 2</i>
pitch	fork
Lawrence	Warden
postal	code
chin	tickled

STEP 2: Determine the proportion of each initial phoneme in each list

<i>List 1</i>	<i>List 2</i>
/p/ = .50	/f/ = .25
/l/ = .25	/w/ = .25
/ç/ = .25	/k/ = .25
	/t/ = .25

STEP 3: Strip each word off its initial consonant

<i>List 1</i>	<i>List 2</i>
-itch	-ork
-awrence	-arden
-ostal	-ode
-in	-ickled

STEP 4: Combine stems from list one with the initial consonants from list two and determine if the resultant strings are words

	/f/	/w/	/k/	/t/	
	.25	.25	.25	.25	sum
-itch	0	1	1	0	.50
-awrence	0	1	0	1	.50
-ostal	0	0	1	0	.25
-in	1	1	1	1	1.00

0 = non-word , 1 = word Average sum = .56

That means the probability that the first outcome of an exchange will create a new word is .56

STEP 5: Combine stems from list two with the initial consonants from list one and determine if the resultant strings are words

	/p/	/l/	/ĉ/	
	.50	.25	.25	sum
-ork	1	0	0	.50
-arden	0	0	0	.00
-ode	0	1	0	.25
-ickled	1	0	0	.50

0 = non-word , 1 = word Average sum = .31

The probability that the second outcome of an exchange will create a new word is .31

Note to the experiment: The decision whether a word is a legitimate word or not is based solely on the real existence of that word in English language. That means, some outcome words sounds phonologically plausible, but won't be a legitimate word in English language. Those outcomes are treated as non-words or non-sensical words.

Figure:2 presents the obtained proportions of words in the various error types and the chance expectation. The Actual Toronto Corpus showed roughly 60% of the first outcomes created words in the case of word exchange. The proportions of words created by exchanges are compared with chance proportions. The results shows that there is a strong lexical bias in the first outcome string and not in the second. This in turn supports the fact that lexicon plays a role in the initial part of the error and therefore, that the phonological stage does access the mental lexicon. These results are consistent with Baars et.al's results and they have a substantial empirical basis, as the experiments were controlled and observations from naturally occurring errors. The results does not support the independence of the positional and functional stages. Errors that seem to be simple phonological misplacements are in fact a results of the accessing of the mental lexicon by the positional stage.

Hypothesis 2 : Functional stage blind to positional stage ?

Lexical selection and lexical insertion are two processes that happens in the functional stage. The lexical item is selected and inserted from the mental lexicon which stores scores of such word forms (Lemma's). Word substitution and word blending errors are typically classified as lexical selection errors. Word substitution error occurs when the intended word is replaced by completely different word. Similarly, word blending occurs when two nearly synonymous (semantically related) words blend. The interacting elements in these errors belong to the same grammatical class or they are semantically related in one way or the other. For example, elbow and knee or taxi and cab. It is very much intuitive to think that, semantically related words gets replaced for one another (elbow and knee). But it is not the case when substitution / misplacement occurs for phonologically similar words, for example present and pressure.

Dell and Reich argues that this phenomenon of phonological influence on word substitutions are due to some kind of "leak" happening between the positional and functional stages. They argue for the activation of nodes in the mental lexicon which results in the misfiring of the wrong words in word substitution errors.

The errors used in this analysis consisted of word substitutions and blends from the Toronto corpus. There were 289 substitutions and 63 blends in total. The phonological similarity of the two words of each error was assessed by determining the percentage of cases in which the two phonemes in the first, second, third and fourth position were identical. The percentages of the the actually observed results as well as the chances are summarized in the

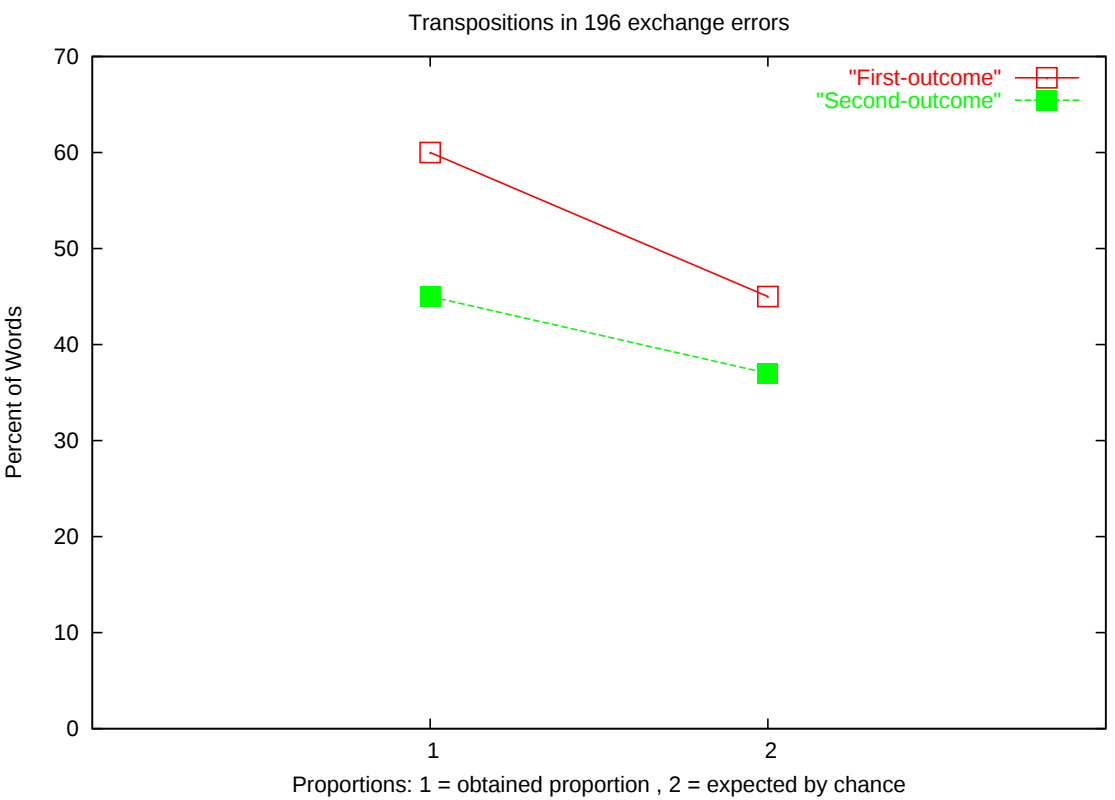


Figure 2. Proportion of word outcomes in initial consonant errors and chance estimates

figure: 3. The numbers shown in the brackets are the number of cases on which each percentage is based. The figure clearly shows a tendency for phonological similarity between the interacting words in both substitutions and blend errors. Therefore, the proposal of leakage between the stages of production holds.

The findings of the Dell and Reich paper can be concluded in terms of some specific findings in terms of the leaks through the leak through the lexicon, as follows

- The lexicon cannot be viewed as a dictionary-like listing of words with separate access files. Instead a network representation consisting of nodes representing linguistic units was postulated.
- Retrieval process in the lexical network occurs by spreading activation of nodes that sends a proportion of its activation to all nodes connecting to it.
- Each such activations also activated phonologically similar nodes.

These findings prove that there are evidences for the hypothesis that at the grammatical stage, the phonological set of nodes are also activated in addition to the semantic nodes. This is one of the reason why phonological similar words are triggered during the speech production. This proposal of spreading activation was later developed into an elaborate theory of interaction between the stages of speech production.

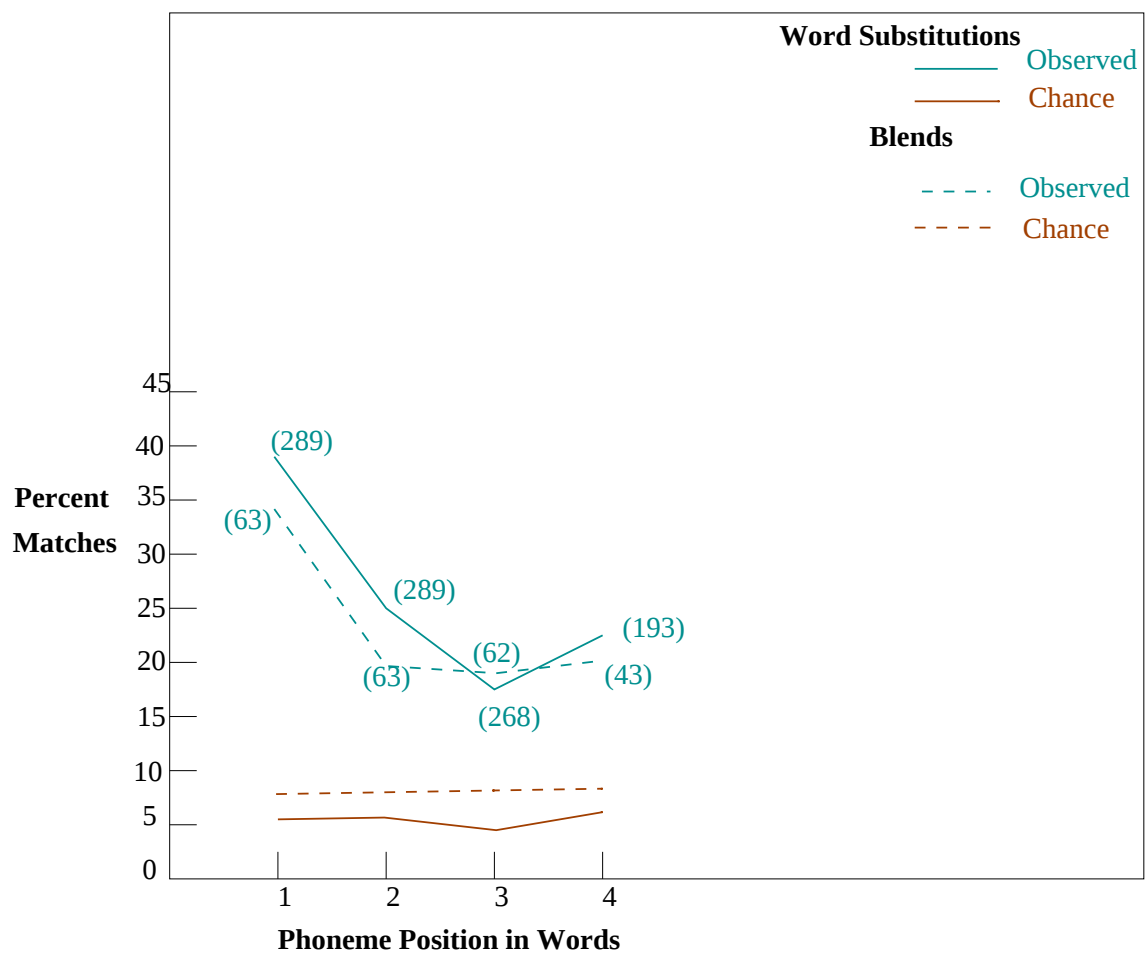


Figure 3: Results of the test for phonological similarity between interacting elements in word substitutions and word blends

5 Gender ascription during tip-of-the-tongue states among Italian speakers

In this section, we will summarize and discuss about the grammatical gender issue on the tip-of-the-tongue (TOT) among Italian speakers, as it is studied in the paper [VAG97]. This research investigates around the question of whether and how syntactic information about words in the mental lexicon. In order to produce syntactically correct expressions, the speakers must have accessed the store of language specific syntactic information. The methodology used for this by Vigliocco et.al was to study the TOT states, which happens when the normal lexical retrieval process breaks down. The TOT states were studied among the Italian speakers, and specifically they were asked to articulate the gender of the words even though they are unable to articulate the word as such. The TOT happens mainly when we fail to recall a word for which we have established knowledge. To utter a word, the speaker must access to the store of lexical representation for providing the lexical semantic and syntactic features of the word and its phonological form. Vigliocco et.al also follows the previous model of language production where the speaker sequentially accesses the distinct representations. Figure : 4 illustrates this model of speech production.

In this figure we can see that the speaker, in order to articulate his/her intention that is the message, firstly transforms this message into a syntactic representation. The construction of this syntactic structure would start with the retrieval of lexical entries (lemmas) whose features are used in the stage of grammatical encoding. This is same as the functional stage as described in the Figure : 1. During this stage, the words are assigned with their grammatical functions (subject, object etc.). This is then inputed to the phonological encoding module, where the appropriate phonological forms are assigned for the words. Here, the lexemes (lexical elements) are retrieved from the store with the correct intonation, accent features etc. This is finally then inputed to articulatory organs which produce overt speech.

The Experiment

As mentioned previously, the TOT phenomenon was studied to examine whether the lemma retrieval fails during the TOT states. Gender markings of the Italian language was chosen for this purpose because, it is purely a linguistic property of the language. That is to say that one cannot distinguish the gender by looking at the endings of the words (as it could be sometimes in German). For example, *sasso masc.* and *pietra (fem.)* both means stone. Here, one cannot make out the gender going by the endings of the

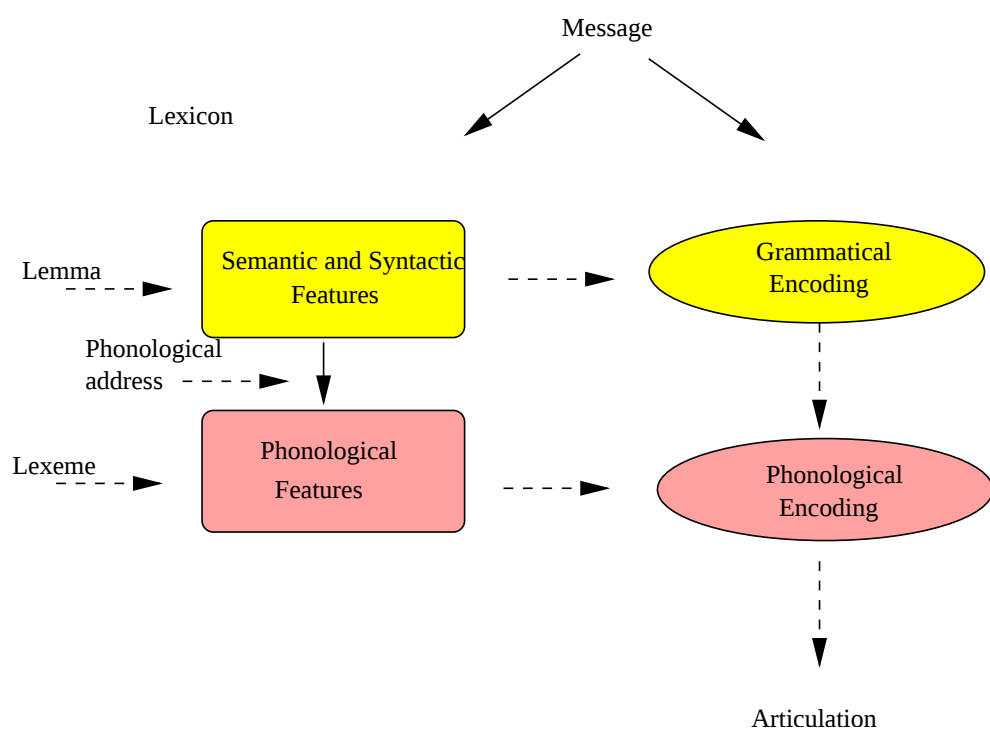


Figure 4: Different levels of representation within the speech production system

words. There were two predictions following the TOT state: The speaker can report the syntactic feature like gender etc. while at the TOT state. They know about the various syntactic features but are unable to report about the phonological information about the word.

In the experiment, 60 subjects were given dictionary like definition of words and asked to guess what the word was. In case of difficulty, they were prompted with the following questions.

- Rate how well they know the word (1=not at all; 5 on TOT state)
- Guess the gender (fem. / masc.)
- Guess the number of syllables
- Guess the letters and their position
- Report any other word that comes to mind

Finally, when there were given the correct word, again they were asked if they knew the word and is the word correct with respect to the expected answer.

The results proved that the Gender was correctly reported in 84% of the cases of TOT states. That means, even though they are unable to pronounce the correct word, they could articulate the gender of the word. Two important claim made through this study was that, the gender is dissociated from the phonological stage of representation. Secondly, this guarantees the independence of grammatical stage and phonological stage. This proves that, even though at the phonological encoding stage the exact information of the word is missing, still it holds the informations of the gender markings of the word which was failed to retrieve from the lexicon.

6 Experiments on Lexical-syntactic Retrieval in TOT States

In this section, we discuss the experimental findings in [MC97] that studies about the retrieval of lexical-syntactic features in the TOT states in Italian. The experiments here replicated and extended the results of Vigliocco et.al [VAG97]. In their studies, Miozzo and Caramazza investigated whether grammatical gender, which is a specific linguistic phenomenon to Italian, could be retrieved by the subjects who find themselves in the TOT states.

In their experiments, the subjects were given pictures and asked to name them. The results shows similarities with the Vigliocco et.al's results but with

some minuscule deviation from the 84% correctness in the gender naming. This deviation was due to the probability of recognizing the correct genders were significantly higher than the "don't know" (DK) responses. Another aspect of the results was that the subjects did not recall the final phoneme better than would be expected by chance and comparable levels of accuracy were obtained for the final phoneme with TOT and DK responses. Further, in none of the experiments the gender could be inferred from the semantic features of the concept (For example, a word having masculine attributed in its meaning).

The statistical results prove that, the grammatical information is specified and accessed independently of meaning and phonological information. Selection of a lexical-semantic representation does not guarantee the access to its grammatical features. Access to phonological features of a word may not strictly depend on access to its grammatical features.

7 Comparison of these experiments and Discussion

In the Dell and Reich experiments (1981), the independence between the phonological (positional) stages and grammatical (positional) stages were compared. In contradiction to one of the previous line of research done by Garret, they proved that there exists dependency between these two stages. That is to say that, the information is to some extent available for both the stages and that they are not completely closed to each other. The studies were done to a convincing extent with the Toronto corpus of errors containing word substitution errors, misplacement errors, blends etc. After Fromkin's corpus of 1973, this was the largest corpus to be collected. The main achievement of the Dell and Reich experiments was the subsequent model of interactive network in speech production.

Later in this report, we dealt with the Vigliocco et.al's result which investigated further in the lines of Dell and Reich's model. In their experiments, they tried to find out how/whether syntactic informations about the words are stored in the mental lexicon which is accessed during the grammatical encoding stage. This was done by examining the tip-of-the-tongue stated in Italian speakers. They found through their experiments that there are evidences that grammatical stage and phonological stages are independent to each other. This refutes Dell and Reich's hypothesis that both the stages have access of information about each other.

Further down in this report we saw the results of Miozzo and Caramazza which was an extension of Vigliocco et.al's paper which supports the interactive stages of processing approach. Their results were in accordance to the

Dell and Reich's hypothesis that the phonological and grammatical stages are not independent modules. This slightly improved and the results of Vigliocco et.al.

8 Conclusion

In this report, we presented mainly three papers. The first paper discussed a seminal result about the stages in sentence production. The basic model Dell and Reich adopted to investigate was Garret's that considered the serial modular flow of information from the functional (grammatical) stage to the positional (phonological) stage.

Nevertheless, one should take these results along with the fact that the "collectors" of the Toronto corpus might have had their perceptual biases making an influence on the collected the data. But, given the fact that, the data were collected by different people each for a duration of one month, and 200 such first hand collections, makes the Toronto corpus fairly reliable as well as the experiments.

There seems to be an agreement among the psycholinguistics researchers about the existence of two stages in the speech production. That is the functional (grammatic) stage where the grammatical and semantic encoding of speech production takes place; and a positional (phonological) stage where the phonological arrangement of sound forms takes place. But at this point, the agreement stops. There exists serious disagreement about the nature of the fundamental issues pertaining to these stages of production.

References

- [DR81] G.S. Dell and P.A. Reich. Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning & Verbal Behaviour*, 20(6):611–629, 1981.
- [MC97] M. Miozzo and A. Caramazza. Retrieval of lexical-syntactic features in tip-of-the-tongue state. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23:1410–1423, 1997.
- [Pot80] John M. Potter. What was the matter with dr. spooner. In Victoria A. Fromkin, editor, *Errors in Linguistic Performance*, pages 13–34. Academic Press, 1980.
- [VAG97] Gabriella Vigliocco, Tiziana Antonini, and Merrill. F. Garrett. Grammatical gender is on the tip of italian tongues. *Psychological Science*, 8(4):314–317, 1997.