CHAPTER II

LITERATURE REVIEW

2.1 Theoretical Framework

2.1.1 Speech Production

Speech production has a strong relationship with slips of the tongue. It has a significant contribution to the understanding of speech processing and language development. One of the most influential psycholinguistic models for speech production has developed by Levelt (1993). There are three stages that Levelt suggested to investigate the production of speech; Conceptualization, Formulation, and Articulation.

The first stage is Conceptualization. This level focuses on what to say and determines the concepts that are to be expressed. During conceptualization, speakers conceive an intention and select relevant information from memory or the environment in preparation for the construction of the intended utterance. For example in the sentence that given by Levelt: "Jhon gave Mary the book" In the speaker's mind, there is a man, the man has a book and he does something to the book (gave it to Mary).

In order to encode a message, the speaker must have access to two kinds of knowledge which is procedural knowledge and descriptive knowledge. In procedural knowledge, the intention of words can deposit the result in *Working Memory*. Working Memory contains all the information currently accessible to the speaker. In

8

other hand, descriptive knowledge is the knowledge that available in *Long-Term Memory*, the speaker's structural knowledge of the world and himself, built up in the course of a lifetime.

The second stage is Formulation. This level involves translating the conceptual representation into a linguistics form. There are two major components to this process: lexical retrieval (selection of the individual words that we want to say) and syntactic planning (putting the selected words together to form a sentence). The next step is putting all of the selections to form a sentence or called syntactic planning. For example in the sentence above, the lexical retrieval shows when the speaker selects "Jhon" as a subject and "the book" as a noun. The speaker places the verb "gave" between the subject and the noun. The next step is putting all of the selections to form a sentence or called syntactic planning.

The final step is to execute the contents of the articulatory program called Articulation stage. This level is phonological level and results in audible sounds. It is involving detailed phonetic and articulatory planning.

2.1.2 Slips of the Tongue

Speech errors are occasionally commented in upon in everyday life. Garman (1990) said that most speakers have had the experience of hearing themselves say something not quite as they intended. Many of these naturally occurring speech errors are either articulatory in nature, as when speech sounds get transposed. Speech errors are errors in spontaneous speech and not the product of intentional ungrammatically or dialects. They occur when the speaker's actual utterance differs in some way from
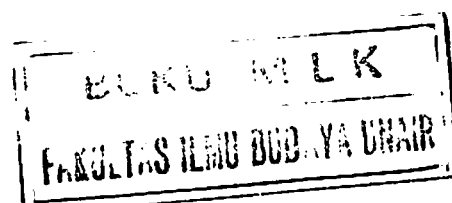
the intended utterance or called the target (Schulze, 2005). Baars (1992) characterized slips of the tongue as errors which are beyond the speakers' control, which are not representative of their ordinary language use, and which can be corrected by them.

What are typically called speech errors are usually associated with the construction of syntactic or phonological representations, and not with the assembly of a motor program. Speech errors involving the misordering or substitution of words (e.g., "writing a mother to my letter") are assigned to the processes that associate lexical items (Baars, 1992).

Garman (1990) used the term 'slips of the tongue' to refer to all types of error manifested in speech production. Slips of the tongue can involve the entire words. A common type of error is a blend of two words: "shrig soufflé" for "shrimp and egg soufflé," "prodeption of speech" for "production and perception of speech".

2.1.3 Types of Slips of the Tongue

Fromkin (1973) classified slips of the tongue based on the error mechanism or underlying units of linguistic performance such as substitution, omitting, transposing, or addition. There are 7 types proposed by Fromkin: (1) Anticipations; (2) Perseveration; (3) Reversals; (4) Blends; (5) Haplologies; (6) Misderivations; and (7) Word Substitution.

In other hand, according to Harley (1995) there are many different types of speech error. He categorizes them by considering the linguistic units involved in the error (for example, at the phonological feature, phoneme, syllable, morpheme, word, phrase, or sentence level) and the error mechanism involved (such as the blend,

substitution, addition, or deletion of units). There are 11 types proposed by Harley:

(1) Phoneme anticipation; (2) Feature perseveration; (3) Phoneme perseveration; (4)

Word blend; (5) Affix deletion; (6) Phoneme deletion; (7) Word substitution; (8)

Word exchange; (9) Morpheme exchange; (10) Phoneme exchange; and (11) Phrase

blend. Here the comparison between Fromkin (1973) and Harley's (1995) types:

Table 2.1 The Comparison between Types of Slips of the Tongue
Proposed by Fromkin (1973) and Harley (1995)

Fromkin (1973)			Harley (1995)		
Type	**Utterance**	**Target**	**Type**	**Utterance**	**Target**
Anticipations	*b*ake my bike	take my bike	Phoneme anticipation	The *m*irst of May	The *first of May*
Perseveration	pulled a *p*antrum	pulled a tantrum	Feature perseveration	Turn the kno*p*	Turn the kno*b*
Reversals	*f*ats and *k*odor	katz and fodor	Phoneme perseveration	God rest *re* merry gentlemen	God rest *ye* merry gentlemen
Blends	*ghast*ly / *grast*ly	grizzly	Word blend	The *chung* of today	The *young children* of today
Haplologies	*Posties*	post toasties	Affix deletion	The chimney catch fire	The chimney catch*es* fire
Misderivations	an interveni*ent* node	an intervening node	Phoneme deletion	Back*gound* lighting	Back*ground* lighting
Word substitutions	before the place *closes*	before the place opens	Word substitution	Get me a *fork*	Get me a *spoon*
			Word exchange	Guess whose *mind* came to *name*?	*Guess whose name* came to *mind*?
			Morpheme exchange	I *randomed* some *sample*	I *sampled* some *randomly*
			Phoneme exchange	Do you *r*eel *f*eally bad?	Do you *f*eel *r*eally bad?
			Phrase blend	Miss you *a* very much	Miss you very much

Based on the types of slips of the tongue above, there are several similar types proposed by Fromkin and Harley which are anticipations, perseveration, reversals, blends, and word substitutions. Fromkin and Harley have different terms to categorize the types. However, they have same main ideas in their explanation. For instance, Harley's phoneme perseveration deals with how a linguistic unit (feature and phoneme) of a word perseverated. That kind of parameter is similar to Fromkin's perseveration. Fromkin's reversals deal with two phonemes of the two words that are exchanged in one utterance. That kind of parameter is similar to Harley's phoneme exchange. Fromkin's anticipation deals with a phoneme that is preceded in one utterance. This kind of parameter is similar to Harley's phoneme anticipation. Fromkin's blends is similar to Harley's word blend because they deal with the two words that blending into one word. Fromkin's word substitutions is similar to Harley's word substitution because they deal with a word that substituted with another word. However, there are some of Fromkin's types which are not included Harley's types of slips of the tongue, which are haplologies and misderivations.

Therefore, the types of slips of the tongue suggested by Fromkin could be supported by Harley's theory for this study. The writer decides to combine the types of slips of the tongue by Harley and Fromkin since both of them are complementary to each other. Here the combination types of slips of the tongue proposed by Fromkin and Harley: (1) Haplologies, (2) Misderivations, (3) Feature perseveration, (4) Phoneme anticipation, (5) Phoneme perseveration, (6) Phoneme exchange, (7) Affix

deletion, (8) Phoneme deletion, (9) Word blend, (10) Word exchange, (11) Morpheme exchange, (12) Word substitution, (13) Phrase blend.

The explanation of each type and examples of that combination as follow:

1. Haplologies

   The elimination of a syllable when two consecutive identical or similar syllables occurr. For examples, post toasties ⟶ posties. Morphology Phonology ⟶ Morphonology.

2. Misderivations

   Misderivations is an error in the process of forming a new word. For example, an intervening node ⟶ an intervenient node.

3. Feature Perseveration

   In feature perseveration, final sound of one word changed yet still in one manner of articulation. For example, *Turn the knop* while the target is *Turn the knob.*

4. Phoneme Anticipation

   In phoneme anticipation, the phoneme in one word changed because the speaker precedes the next following word.

   For example, *The mirst of May* while the target is *The first of May.*

   That mook is mine while the target is *that book is mine.*

5. Phoneme Preservation

In phoneme preservation, the phoneme in one word changed because the speaker preserves the previous word. For example, *God rest re merry gentlemen* while the target is *God rest ye merry gentlemen.*

6. Phoneme Exchange

In phoneme exchange, the phonemes of two words are exchanging. Usually it happens because of the similar vowel between the two words. For example, *Do you reel feally bad?* while the target is *Do you feel really bad?*

I neally reed this job while the target is *I really need this job.*

7. Affix Deletion

In affix deletion, the affix of the word is deleting. For example, *The chimney catch fire* while the target is *The chimney catches fire.*

Keep fighting! nothing is possible while the target is *nothing is impossible.*

Affix means a morpheme that is attached to a word stem to form a new word, for instance, prefix and suffix. Prefix is an affix that appears before a stem (un-, anti-, auto-, im-, hyper-, intra-, and etcetera) while suffix is an affix that appears after a stem (-ing, -s, -ed, -t, -en, -er, est, -n't) suggested by Plag and Braun (2007)

8. Phoneme Deletion

In phoneme deletion, one phoneme of the word is deleting. For example, *Backgound lighting* while the target is *Background lighting.*

Band strategy while the target is *brand strategy.*

9. Word Blend

In word blend, two words are blending into one word. For example, *The chung of today* while the target is *The children young of today*.

Give me the spork! while the target is *give me the spoon and fork!*

10. Word Exchange

In word exchange, two or more words in one phrase are exchanging. For example, *Guess whose mind came to name?* while the target is *Guess whose name came to mind?*.

The layer burns the fire while the target is *the fire burns the layer*.

11. Morpheme Exchange

In morpheme exchange, two morphemes in one phrase are exchanging.

For example, *I randomed some samply* while the target is *I sampled some randomly*.

The water was boilly amazing while the target is *the water was amazingly boiling*.

Morpheme is the smallest unit of meaning in a language. Morphemes that are also words are called free morphemes, suggested by Carroll (2008).

12. Word Substitution

In word substitution, the word in one phrase is substituted by another word.

For example, *Get me a fork* while the target is *Get me a spoon*.

Could you please bring my pen? while the target is *could you please bring my pencil?*

13. Phrase Blend

In phrase blend, two phrases are blending in one phrase. For example, *Miss you a very much.* It could be *miss you a great deal* that is blended with *miss you very much.*

Thank you a bunch. It could be *thank you very much* that is blended with *thanks a bunch.*

2.1.4 Indonesian Affixes

In Indonesian language affix means *"imbuhan"*, it consists of prefix or *"awalan,"* infix or *"sisipan,"* suffix or *"akhiran,"* and circumfix or *"imbuhan awalan dan akhiran."* (Alwi, Dardjowidjojo, Lapoliwa, & Moeliono, 2003).

Prefix is an affix which attached before the root of a word (e.g. *di-* in *makan* ⟶ *dimakan*). There are seven prefix in Indonesian language. The prefixes in Indonesian language are *ber-, di-, ke-, meN-, (me, mem-, men-, meng-, menge-, meny-), peN- (pe-, pem-, pen-, peng, penge-, peny-, per-), se-,* and *ter-.* The letter *N* in prefix *meN-* and *peN-* is modifiable; it becomes Nasal phoneme depending on the first letter of the root word. Some of Nasal phonemes are /n/, /m/, /ny/, and /ng/.

Infix is an affix which inserted inside a root word. There are three infix in Indonesian language. They are *-el-, -em-,* and *-er-.* An infix is usually placed into the root word after the first syllable of the root word (e.g. *-em-* in *pimpin* ⟶ *pemimpin*)

Suffix is an affix which attached after the root word *(e.g. –kan* in *pejam* ⟶ *pejamkan)*. There are five suffix in Indonesian language. They are *-kan, -i, -an, -lah,* and *-kah.* Two suffixes cannot be attached on the same root.

Circumfix is a combination between prefix and suffix (e.g *meN-kan* in *lapor* → *melaporkan, ber-kan* in *baik* → *berbaikan)*. There are fourteen circumfix in Indonesian language; *meN-kan, ber-kan, ber-an, peN-kan, peN-an, per-an, memper-kan, diper-kan, ter-kan, ter-i, di-kan, di-i, ke-an,* and *se-nya*.

2.1.5 Correction in the Slips of the Tongue

The speakers do a correction because they realize that they make a mistake and try to repair their mistake by giving a correct word. The speakers replace certain old words with the new ones, indicating that the old words had been misplanned. Dubois (1975, in Clark 1977) stated that the correction phases 'that is': reference editing, 'or rather': nuance editing, 'I mean': mistake editing, 'well': claim editing. Correction is symbolized by [—]. Below the examples of correction:

- Please, give me the pen—I mean, the pencil!

- He went to the cinema last night—or rather, went to the concert

- Mira loves the ice cream—that is, Lisa does

- I have an appointment with Jack tomorrow—well, on Wednesday afternoon

However, when the slips are undetected by the speaker, it may be caused by several reasons such as the speakers are competing demands for attention, under the influence of alcohol or drugs, in the case of a marked degree of fatigue, or under the influence of psychological stress such as embarrassment, nervousness, anger, fear, et cetera (Laver 1969, in Fromkin 1973).

In order to get the comprehension of the listeners, make a correction is necessary for the speakers after they produced an error in their speech. Therefore, a
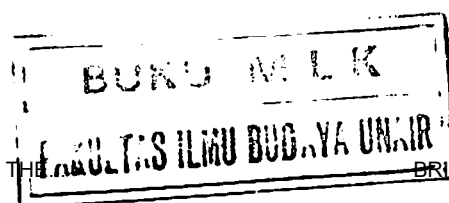
good speaker is always makes a correction. Wijaya (2009) agreed that it is important to note that the announcers should avoid doing nothing after they made a mistake. The good announcers should make a correction immediately when they made a mistake.

2.1.6 Speech Rate

Announcers should have a capability in speech performance. This skill is needed not only for announcers but also for reporters, presenters, and other performers. Speech rate is how fast or slow a person speaks. Everyone has a different rate of speech depending on her/his location, age, culture, and how he/she feels. In order to communicate effectively we must speak at a rate of speech that our listeners can understand. Based on Kendall (2013) speech rate is one of components of linguistic performance rather than linguistic competence. Speech rate is associated with the time of speech. The differences in speech timing are based on a number of social factors. Announcers usually show their skill by presenting their high speech rate. Faster speech is typically associated with competence, intelligence, expertise, social attractiveness (Street et al. 1983, in Kendall 2013), and greater persuasiveness (Miller et al. 1976, in Kendall 2013) over slower speech.

Speech rates vary tremendously among normal speakers; it is difficult to assign a standard word-per-minute (WPM). Adults produce an average of 125 words per minute during conversational speech, and 100-120 words per minute during oral reading. To determine speech rate, tape-record a sample of connected speech (devoid

of significant pausing) in oral reading, in conversational speech, or both. In a 60-second interval, count the number of words produced and divided by 60. For example, 200 words produced in 60 seconds is 200 WPM. If there are no 60-second intervals of connected speech, the procedures as follow:

1. Time the sample (e.g., 20 seconds).

2. Count the number of words produced (e.g., 62 words).

3. Divide the number of seconds in a minute (60) by the number of seconds in sample (20 seconds in the example): $60 \div 2 = 3$

4. Multiply the number of words in the sample (62 in the example) by te number in Step 3 (3 in the example): $62 \times 3 = 186$. The WPM is 186 (Calvert & Silverman 1983, in Shipley & McAfee 2009).

2.1.7 Emotional Condition of the Speaker

Scherer et al. (2001) stated that emotion affects the psychology of speech production. The human voice provides the carrier signal for speech, that is, the sound that is modulated by specific features of language. In situations where the body is in a relatively stable, unchallenged situation, the parts of the anatomy can be relatively freely produced the service of speech production. In other hand, when the speaker gets nervous, unstable emotional condition, it will probably interfere the production of speech. In addition, Carroll (2008) purposed that errors are more common when we are nervous or under stress, as when performers appear on live television and

radio shows; programs devoted to television's best "bloopers" never seem to run out of material. It seems probable that errors are more likely to occur when we are tired, anxious, or drunk.

Juslin & Scherer (2008) argued that speech emotion analysis refers to the use of various methods to analyze vocal behavior as a marker of affect (e.g., emotions, moods, and stress), focusing on the nonverbal aspects of speech. The basic assumption is that there is a set of objectively measurable voice parameters that reflect the affective state a person is currently experiencing (or expressing for strategic purposes in social interaction). Emotion in speech may be regarded as a communication system featuring several parts:

1. the expression or portrayal of the emotion by the speaker (the encoding),

2. the acoustic cues (e.g., sound intensity) that convey the felt or intended emotion

3. the proximal perception of these cues by the perceiver (e.g., perceived loudness)

4. the inference about the expressed emotion by the perceiver (the decoding)

Based on Experiencing speech anxiety is normal. Nearly everyone gets nervous when they have to give a speech or a presentation, even experienced speakers. The speakers that look relaxed and confident have simply learned how to handle their anxiety and use it to enhance their performance. Most of an anxiety is not

visible to the audience. We may feel like we are shaking uncontrollably but people in the audience probably cannot even tell.

Anxiety decreases as a speech progresses. Speech anxiety is usually worst right before a speech and at the beginning of the speech. Most people find that once they get through the introduction their anxiety begins to decrease and confidence increases (University of Pittsburgh, 2008).

In addition, Carroll (2008) stated that errors are more common when we are nervous or under stress, as when performers appear on live television and radio shows; programs devoted to television's best "bloopers" never seem to run out of material. It seems probable that errors are more likely to occur when we are tired, anxious, or drunk.
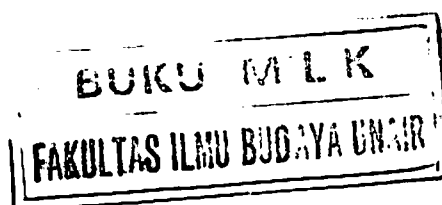
2.2 Review of Related Studies

Some of studies have been done by researchers who are interested in speech errors specifically in slips of the tongue. One of them Kawachi (2002) who has conducted a research *"Practice Effects on Speech Production Planning: Evidence from Slips of the Tongue in Spontaneous vs. Preplanned Speech in Japanese"*. He focused on how practice influences speech production planning processes. Most models of speech production planning have been developed out of errors in speech production planning, specifically "slips of the tongue" (hereafter, SOTs). He compared slips of the tongue in Japanese collected spontaneous everyday conversation and those collected from largely preplanned conversation in live-

broadcast TV programs. The researcher collected slips of the tongue from two types of sources, spontaneous, face-to-face everyday conversation, and videotaped, live broadcast TV programs, mainly talk shows and entertainment shows, where speech was presumed to have been preplanned to a large extent and no reading was involved. From the result he found a variety of practice effects on speech production planning by comparing slips of the tongue in everyday conversation and those in TV programs, although this study has also shown that there are some aspects of the speech production planning mechanism that are insusceptible to practice.

In addition, another study discussing about speech errors was written by Gordon (2007), he conducted a study *"Interpreting Speech Errors in Aphasia"*. He discussed the distributions of word-retrieval errors produced by subjects with aphasia. He used thirty-two subjects with aphasia, representing a range of aphasia subtypes and severity levels, participated in the study. They named 175 line drawings of objects from the Philadelphia. The result of his study showed the distribution of semantic errors is unimodal, peaking at 6%, with one outlier, indicating that the occurrence of semantic errors is unlikely to reflect the presence or absence of an underlying semantic impairment. The distribution of phonological errors, on the other hand, is bimodal, with one peak at 2% and another at 16%, suggesting that most aphasic speakers produce a few phonological paraphasias, but some produce an abundance of them.

The next related study comes from Nisa (2009) who has conducted *"A Study on Speech Errors Made by Global TV VJ the Music Programs: Most Wanted"*. She described about speech errors that occur on the presenter as known as VJ of Global TV music program. She focused on the common types of speech errors such as pauses, repeats, and false starts. She collects the data from VJs Global TV when she/he presenting the music programs. The result showed that hesitations (filled pauses and unfilled pauses) are the most frequent speech errors that are uttered by VJs (42,52% and 29,35%). The next following speech errors are repeats (9,58%), corrections (4,75%), retraced false starts (4,20%), stutters (4,20%), interjections (3,00%), unretraced false starts (1,80%), and slips of the tongue (0,60%).

The first related study above discussing about slips of the tongue by Kawachi. He focused on the effect of practice to the speech production planning processes. It supported further for the writer's study. However, it was totally different since the writer focused on the types of slips of the tongue itself which produced by radio announcer. Even tough, the second and the last third studies above focused on the speech error, yet it was also different with the writer's study. Gordon interpreting speech error in Aphasia and Nisa described types of disfluency in speech production while the writer focused on the types of slips of the tongue by Fromkin and Harley's theory.

CHAPTER III
METHOD OF THE STUDY