The lexical bias effect in experimentally elicited Swedish phonological speech errors

Hanna Roskvist

Lund University, Centre for Languages and Literature

BA Thesis in General Linguistics, ALSK13

Supervisor: Sandra Debreslioska

8/5-2025

Abstract

The lexical bias effect refers to the tendency for phonological speech errors to have lexical rather than non-lexical outcomes. This effect has often been used as a piece of evidence in theories of speech production, informing assumptions about monitoring and spreading activation. However, the lexical bias effect is known only from investigations of a relatively small set of languages, and the findings have not always been conclusive. The aim of the current study is to investigate the presence of a lexical bias effect in Swedish phonological speech errors, which has not been done before. Errors were elicited by using a modified version of the SLIP-method (Baars et al. 1975). The errors elicited in the current study display a strong lexical bias, further indicating that the effect is universal and not language-dependent. Furthermore, the errors displayed a phonetic similarity effect as well as a preference towards grammatical and meaningful error outcomes, revealing additional factors underlying speech errors and speech production in general.

Content

1. Introduction

The terms ‘speech error’, ‘slip of the tongue’, or sometimes ‘Freudian slip’ denote the various unintentional mistakes which we all make when speaking, and which constitute a normal part of everyday language use. For example, personal observations[1] of speech errors include utterances like ‘There is a limit to how much beer you can drink’ →[2] ‘There is a limit to how much drink you can beer’, ‘the Russian president’ → ‘the Russident’, and ‘I will briefly boil it’ → ‘I will beefly broil it’. While speakers often find these errors humorous and sometimes embarrassing, errors tend to be promptly brushed off and forgotten. They are, after all, a distraction from what we are actually trying to communicate. However, though speech errors might be unintentional and unwanted, they are by no means worthless to the psycholinguist. Speech errors have long been taken as evidence for the universal processes of speech production in general, because errors are highly regular and seem to strictly adhere to various patterns which reveal how they were constructed in the mind (e.g., Fromkin, 1973; Nooteboom, 1973; Garrett, 1975; Dell & Reich, 1981). It must be assumed that non-erroneous speech adheres to the same patterns that errors do, which allows speech errors to open a window into how the human mind organizes linguistic components when putting together an utterance.

One such revealing pattern or tendency within speech errors is the so-called lexical bias effect. The term refers to the tendency for phonological speech errors (i.e., errors including the misplacement of sounds, like ‘briefly boil’ → ‘beefly broil’) to have lexical rather than non-lexical outcomes. This means that phonological errors are more likely to occur if the resulting words are real words than if they are not. In other words, an error like ‘darn bore’ → ‘barn door’ is more likely than the error ‘dart board’ → ‘*bart *doard’, because the former outcome is lexical and the latter is non-lexical (e.g., Baars et al., 1975; Dell & Reich, 1981). Since the 1970s, many researchers have been interested in the lexical bias effect as a piece of evidence for processes of speech production. The effect has been used to inform key theoretical assumptions about how these processes take place, for example in discussions on monitoring and modularity in speech production. The lexical bias effect is currently fairly well-attested in experimental research, but this research has been conducted on a relatively small number of

[1] The errors from “personal observation” have been collected by myself in everyday settings over a long period of time. Most of these errors cited in this thesis were produced not by L1 speakers of English, but by highly proficient L2 speakers (L1 being some other Germanic or otherwise Indo-European language) or early bilinguals. Evidence suggests that speech errors produced by L2 speakers and by L1 speakers are more similar than not – see Poulisse (1999) for a comprehensive overview of this topic.

[2] Transformation arrows are conventionally used when describing a speech error – to the left of the arrow is the utterance that was intended, and to the right of the arrow is the erroneous utterance that was actually produced.

languages, namely English (Baars et al., 1975; Motley et al. 1982; Dell, 1986), Dutch (Nooteboom, 2005; Nooteboom & Quené, 2008; Nooteboom & Quené, 2017), and Spanish (Hartsuiker et al., 2006; Costa et al., 2006).

In order to base assumptions about universal processes of speech production on the lexical bias effect, it is important that the effect is indicated to hold cross-linguistically. Therefore, the study of the lexical bias effect would benefit from being expanded to include more languages. There have as yet been no studies on the lexical bias effect in Swedish. In order to fill this research gap, the aim of the current study is to investigate the presence of a lexical bias effect in Swedish phonological speech errors. This will provide increased clarity as to how well the effect holds cross-linguistically. This aim was achieved by experimentally eliciting phonological speech errors and investigating the degree of lexical bias in these.

2. Background

I will begin this section with a general definition of speech errors as a phenomenon and as a field of study. Secondly, I will give an overview of speech production, introducing one model of speech production, and how speech errors can be understood within this model. Thereafter, I will describe the lexical bias effect and how it has been related to theoretical assumptions about speech production. Finally, I will introduce the SLIP-method, which is an experimental approach to eliciting speech errors commonly used to investigate the lexical bias.

2.1 Speech errors

A 'speech error', also called 'slip of the tongue' or simply 'slip', is defined as "an involuntary deviation in performance from the speakers current phonological, grammatical or lexical intention" (Boomer & Laver, 1973, p. 123). Such an involuntary deviation can originate at any stage during speech production, which is evidenced by the existence of higher-order errors like semantic errors (e.g., 'I'm gonna go to bed' → 'I'm gonna go to night', from personal observation). This shows that speech errors are more than purely articulatory mistakes, which the term 'slip of the tongue' might otherwise imply. For this reason, the current thesis favors the term 'speech error' over 'slip of the tongue'. To reiterate, a would-be speaker has a phonological, grammatical or lexical intention, that is, a target utterance, that they wish to realize. The defining trait of a speech error is that it misses this target in some way. A speech error is an error not because it is (necessarily) ungrammatical or nonsensical, but because it differs from the speaker's target.

The present thesis is concerned only with phonological speech errors. Simply put, phonological speech errors are errors involving the misplacement of some phonological element(s), such as features, phonemes, or syllables. Such misplacements can result in errors like exchanges, anticipations, and perseverations. An exchange is when two elements swap places with each other, so that one intrudes into the slot of the other and vice versa. For instance, a spoonerism is a certain type of exchange where two word onsets are exchanged with each other, for example 'featured picture' → 'peatured ficture' (personal observation). An anticipation error is when a segment appears earlier than it was supposed to, for example 'very briefly' → 'very bliefly' (personal observation). A perseveration is the opposite, that is, when a segment appears later than it was supposed to, for example 'a phonological rule' → 'a phonological fool' (example from Fromkin, 1973). These three types of phonological speech errors are those that will be elicited in the current study to investigate a lexical bias effect and underlying speech production processes.

2.1.1 The significance of speech errors

While speech errors have received attention partly by merit of being interesting per se, they are a staple of psycholinguistic research because they pertain to speech production more generally. It has long been known, mostly from studying corpora of naturalistic speech errors, that errors tend to occur in highly regular ways and follow certain patterns (e.g., Fromkin, 1971; Garrett, 1975; Dell & Reich, 1981). Since normal speech and erroneous speech must both be assumed to be the result of the same speech production processes (i.e., proceeding normally in the former case and abnormally in the latter), the patterns found in speech errors must be assumed to apply to normal speech as well. This makes speech errors useful, because it is not necessarily obvious from a normal utterance how the underlying linguistic units are structured in the mind. An utterance like 'I see the cat' does not tell us, say, whether or not the linguistic categories 'verb' and 'noun' are psychologically real, that is, whether or not they are treated as different categories in the mind. Speech error data, however, can reveal underlying processes through their patterns: For example, the unit similarity constraint is the tendency for any given linguistic unit to be more likely to be part of an error with a unit of the same linguistic category or otherwise of similar quality (e.g., Fromkin, 1971). This means that nouns are more easily confused with other nouns, while verbs are more easily confused with other verbs, syllable onsets with other onsets, and so on. That is, saying 'I see the cat' while actually meaning to say 'I see the horse' is a more likely error than 'I cat the horse', showing that nouns and verbs do not behave in the same way. This constraint is just one example in a long array of such regularities that have been observed and used to inform assumptions about underlying speech production processes. Such patterns allow us to infer, for example, that the categories 'verb' and 'noun' are, in fact, psychologically real – otherwise, there should be no tendency for them to pattern as such in the error data. Another instance of the unit similarity constraint is the tendency for a phonological unit to be part of an error with another phonological unit that occupies the same syllable position. That is, onsets are more easily exchanged with other onsets, codas with other codas, and so on. This is illustrated in the case of spoonerisms, which involve the exchange of two word onsets. Again, this tendency serves as another window into how the mind organizes linguistic information: since speech sounds typically exchange with other sounds in the same syllable position, it appears that the syllable is a psychologically real unit in the production of speech.

2.2 Speech production

2.2.1 Activation

An auxiliary concept that is central to the field of speech production, and which must be introduced before describing the processes of speech production in any detail, is activation (see, e.g., Harley, 2014, Appendix, for a thorough explanation of this concept). Activation is a frequent way of conceptualizing the inner workings of any processing component in language use. A speech production system is often thought of as something akin to a neural network, consisting of units which are concatenated by different connections. Each unit (corresponding to, for example, a lexical item or a phoneme) has different degrees of activation that change constantly over time. Activation passes from one item to another through their connections, which have different strength. The strength of a connection between two units is determined by how similar or relevant the items are to each other (for example, the word 'cat' would be more strongly connected to the phoneme /k/ than to /f/, and the phoneme /k/ more strongly connected to a similar sound like /g/ than to a less similar sound like /a/). Typically, the speech production system is conceived of as having to make selections as to what linguistic output it will produce, in choosing the appropriate words to express an idea, and then picking out the appropriate phonemes to produce those words. The factor governing such selections can be understood as different levels of activation assigned to different candidates. A unit's level of activation is an abstract measure of appropriateness for whatever choice the processing component is dealing with – the more appropriate a given item is for a given choice, the higher the activation will be, and once the activation is high enough for a given item, that item is selected. When the system makes a selection, multiple candidates will have received at least some degree of activation (for example, when constructing the utterance 'cat', both /g/ and /k/ might have been activated due to their phonetic similarity), but the candidate with the highest degree wins. The concept of activation will assist in understanding how speech production progresses through different stages and deals with different levels of linguistic units.

2.2.2 Levelt's model of speech production

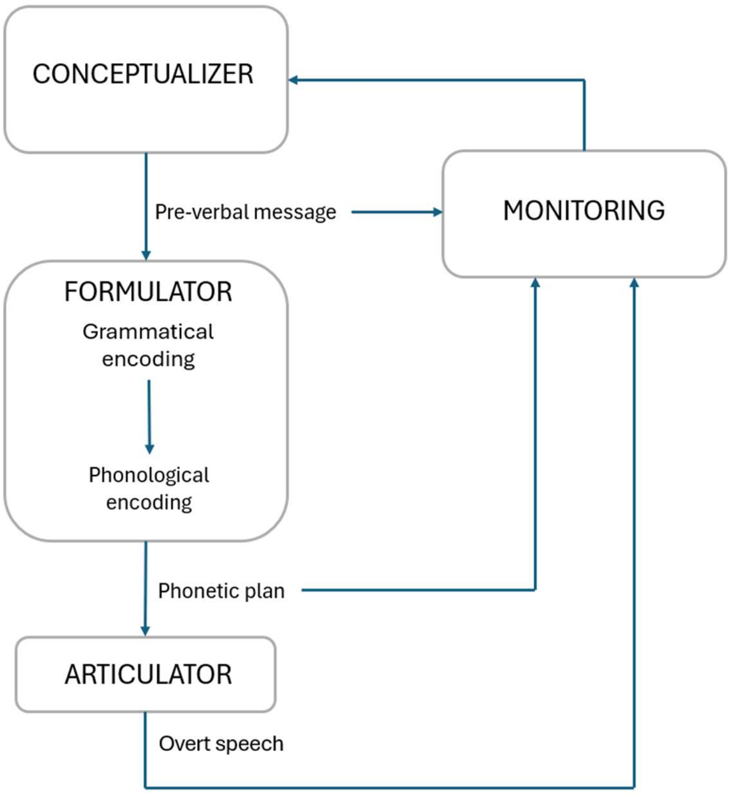
In his prominent model of speech production, Levelt (1989) conceives of speaking as a process which comes to pass through three main processing components: a Conceptualizer, a Formulator, and an Articulator, operating in that order (see Figure 1). All of these components specialize in producing a certain kind of output, which serves as the input for the next component in line. The Conceptualizer is responsible for taking a thought or idea and ordering it into a pre-verbal (i.e., purely conceptual and not yet linguistic) message, which constitutes the output passed on to the Formulator. The Formulator is responsible for filling the pre-verbal message with all necessary pieces of linguistic content. The Articulator is responsible for articulating (i.e., physically executing) an articulatory plan that it receives as input from the Formulator. Finally, an additional component in Levelt's model is monitoring. Monitoring is a process whereby speakers check their own output as well as planned output (i.e., both overt and covert speech) for deviations from the intended message, so that these may be caught and corrected. Monitoring employs the same speech-comprehension system that is used when processing the speech of others, and it runs concurrently to all stages of production. If there are mistakes at any point during formulation so that unintended output is prepared for production, the monitor will catch many of these mistakes and inhibit them before they are articulated.

It is mainly the Formulator which is interesting for the study of phonological speech errors, and will therefore be treated in greater detail here. The Formulator is divided into two main subcomponents: a Grammatical Encoder and a Phonological Encoder. The Grammatical Encoder receives the pre-verbal message from the Conceptualizer and starts mapping its concepts onto specific lexical items. All lexical items are, according to Levelt, twofold: any given lexical item consists of a lemma on the one hand, which contains semantic and syntactic information, and the item's lexical form or lexeme on the other, which contains phonological information. For example, the lemma of the lexical item 'see' contains information such as that it is a verb, that it is transitive, and that it ought to be placed after a subject in the nominative case. The lexeme of the word 'see' contains information such as that it begins with the sound /s/. The Grammatical Encoder selects the appropriate lemmas, arranges them into a syntactic plan called a surface structure, and passes this output on to the Phonological Encoder.

The task of the Phonological Encoder is to fill the surface structure with actual phonological information, so that it can generate an articulatory plan for the Articulator. This is done by accessing the lexemes which correspond to the lemmas and breaking down their phonological makeup into smaller and smaller units that can be specified for articulation. This is done through what Levelt calls spellout procedures. During so-called segmental spellout, the appropriate morphemes for a planned utterance are retrieved and placed in the correct order, and their syllabic and segmental composition is generated (or 'spelled out'). This information about syllabic and segmental composition, that is, information about the syllables and phonemes that should be part of the utterance, is then passed on to the stage of phonetic spellout. Here, the segments are retrieved and placed in the correct order. After this, a phonetic or articulatory plan consisting of specified phones (or, more concretely, sets of articulatory gestures) is generated (see Figure 2). This phonetic plan is then passed to the Articulator to be physically executed.

Figure 1

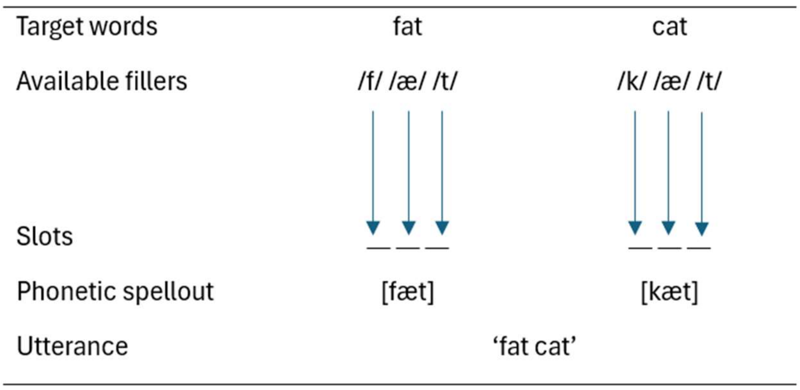
Levelt's model of speech production (Adapted from Levelt, 1989)



Levelt (1989) describes the inner workings of these spellout procedures using a 'slots-and-fillers' framework, particularly inspired by Shattuck-Hufnagel's (1979) 'scan-and-copy' mechanism. The scan-and-copy mechanism can be simplified as follows: during grammatical and phonological encoding of a message, the speaker builds a frame consisting of slots. Into each of these slots, a filler must be inserted. A 'scanner' scans the set of available fillers and copies them into the slots in serial order, meaning 'left to right' from the beginning to the end of a planned utterance. Two steps that govern the insertion of fillers are selection and checkoff: The most activated element is selected to fill a slot, after which it is checked off, not to be used again. This framework can be applied to Levelt's spellout procedures in the following way: Segmental spellout generates the phonemic composition of a word, which is passed on to the stage of phonetic spellout as a frame. During phonetic spellout, the scanner scans the sets of available phonemes, which are then copied to fill the appropriate slots in this frame. After all the slots in the frame have been filled, an articulatory plan is generated (see Figure 2).

Figure 2

Phonetic spellout



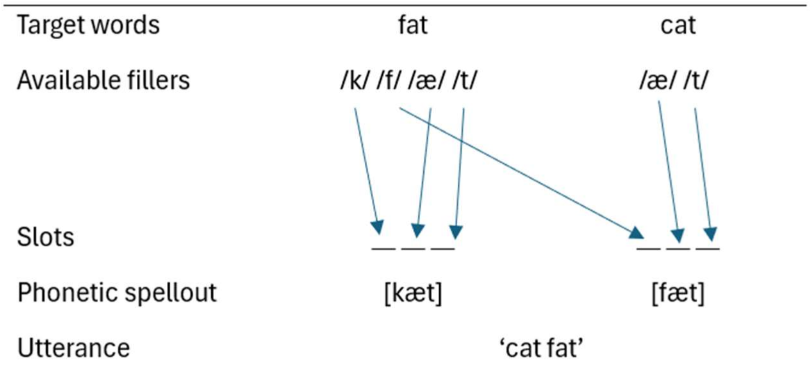
2.2.3 Speech errors in Levelt's model

Most of the time, all fillers end up in the correct slots, and the utterance is produced as intended – but if, at any point, a slot is filled by an inappropriate filler, then a faulty articulatory plan can be passed on to the Articulator and a phonological speech error might be produced. Such a misplacement of a filler can be due to failure in the scan-copier, as well as in the checkoff mechanism: The scanner might make a filler available earlier than intended or later than intended, or make available a filler that is not appropriate for the context at all. The copier might then place that filler into a slot where it was not intended to go, after which it can be inappropriately checked off or failed to be checked off.

For example, when an exchange occurs, two fillers in the form of consonantal onsets might be made available simultaneously, because segmental spellout has progressed quickly. Let us say that the speaker is preparing to produce the words ‘fat cat’. The first consonant /f/ has been scanned and become available for copying, but so has /k/, despite not being required yet. The copier starts filling in the slots from left to right, and must first select whether to put /f/ or /k/, which are both available and appropriate onsets, into the onset slot of the first word. Because the speech production system does not function perfectly all the time, /k/ accidentally happens to receive higher activation than /f/. Consequently, /k/ is selected to fill the first onset slot of the first word, after which it is checked off, meaning that it cannot be used again. When the copier later arrives at the first onset slot in the second word, the correct filler /k/ is no longer available for selection, and because of the unit similarity constraint, it is unlikely that the coda /t/ will be selected to fill an onset slot. Consequently, the only appropriate filler that is still available to put in this slot is /f/, yielding the outcome ‘cat fat’ instead of the intended ‘fat cat’. Figure 3 illustrates such an exchange error.

Figure 3

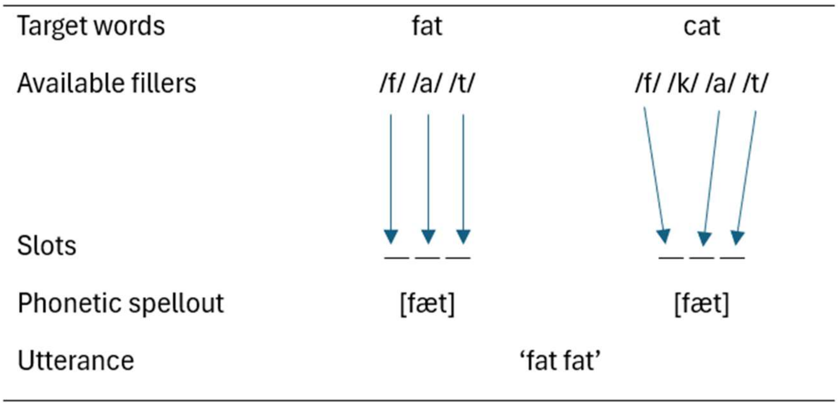
Phonetic spellout: Exchange error



The processes behind anticipations and perseverations are similar: to take the example ‘briefly’ → ‘bliefly’, the two potential fillers /l/ and /r/ are made available at the same time. Another possible scenario is that the inappropriate filler /l/ is made available too early while the correct filler /r/ is made available too late. Regardless, the inappropriate filler /l/ is available and inadvertently receives more activation than the appropriate filler /r/. It is therefore incorrectly selected to fill the second slot in the first onset cluster in ‘briefly’. This time, however, /l/ is not properly checked off after being selected, and it can appear again in its correct slot in the second onset cluster, yielding ‘bliefly’. As for a perseveration error, we can exemplify with the imaginary error 'fat cat' → 'fat fat'. In this case, /f/ is correctly selected to fill the onset slot of the first word, but is not properly checked off. Since it has not been checked off, it is still available when the copier must select a filler for the onset of the second word. If /f/ receives higher activation than the appropriate filler /k/, it will be selected for the second onset slot as well, yielding 'fat fat'. See Figure 4 for an illustration of a perseveration error.

Figure 4

Phonetic spellout: perseveration error



2.2.4 Modularity and interactivity in speech production

There is an important matter concerning the nature of processing in speech production that has been up for discussion and must be mentioned here, in order to facilitate a deeper understanding of speech errors and the lexical bias effect. A long-standing issue in models of speech production and psycholinguistics in general is the question of modularity, that is, whether the different processing components involved in language use are discrete or interactive (see Vigliocco & Hartsuiker, 2002, and Kerr et al., 2023, for summaries and reviews of different claims made on this topic).

The discrete view, simply put, posits that the components in speech production are autonomous entities that produce their specialized kind of output using only the input from the previous component, without requiring or transmitting any superfluous information. Levelt's (1989) model notably assumes such modularity to a high degree. The first component in line must finish producing its output before the next component begins producing its own, and the output that a component passes on is only what is necessary for the next component in line to produce its own kind of output. This means that a component only receives information about selections made at previous levels, and items that received some activation but were ultimately not selected do not pass on their activation to connected items on subsequent levels. Furthermore, a modular view of speech production favors a unidirectional flow of information and activation. That is, a given component only receives information about selections from the previous level, and the activation of items at one level do not affect the activation of items at the previous level (Vigliocco & Hartsuiker, 2002). While a component is encoding, activation does not spread forwards or backwards from one component to another. Consequently, a component cannot update its output according to events at a subsequent level.

The interactive view, on the other hand, considers the production components part of a more interconnected network, in which there is interaction between components at all levels of encoding. In models assuming such interactivity, which can be referred to as spreading or cascading activation (e.g., Dell, 1986), activation passes from one component to the next even before selection is finished. This means that one stage of processing can commence even before it has complete information from the previous one, and that even unselected items that received some degree of activation can affect the activation of items on subsequent levels. The interactive view also favors a bidirectional flow of information, meaning that each component operates using both a forwards and a backwards flow of information and activation (Vigliocco & Hartsuiker, 2002). In this way, each processing component can adapt its own output according to activation fed back from the component in front.

The degree of interactivity in the different connections between different components in speech production is ultimately, as Levelt (1989) points out, an empirical issue. One type of evidence that is used to inform theories on this issue, like for issues in speech production in general, comes from speech errors. The lexical bias effect is one such piece of evidence that has been used extensively in this discussion.

2.3 The lexical bias effect

As mentioned, speech errors are characterized by various regularities that they adhere to more or less strictly. One such regularity is the lexical bias, which is the tendency for phonological speech errors to have lexical rather than non-lexical outcomes. In other words, a speaker is more likely, upon making a speech error, to accidentally produce a real word than a non-word (i.e., a non-existent and phonotactically illegal word) or a pseudoword (i.e., a non-existent but phonotactically legal word). For example, a spoonerism like ‘darn bore’ → ‘barn door’ is more likely to occur than a spoonerism like ‘dart board’ → ‘*bart *doard’. The effect has been found in experimentally elicited errors (e.g., Baars et al., 1975) as well as in corpora of naturalistic errors (e.g., Dell & Reich, 1981). It must be noted, however, that some corpus studies have failed to reveal a lexical bias (e.g., for English in Garrett, 1975; for Spanish in Del Viso et al. 1991).

The lexical bias in speech errors has been considered to reveal underlying processes of speech production. Specifically, it implies that the phonological stage of speech production has access to information about the lexicality of its own output. One might assume that since the Phonological Encoder specializes in producing pronounceable and phonotactically permissible sequences, it is automatically biased towards lexicality over non-lexicality. However, there is a practically infinite amount of such permissible sequences in any given language that are *not* lexical (i.e., pseudowords), and such sequences are also dispreferred in favor of lexical sequences. If lexical status was unimportant during phonological encoding, we should not expect to find any lexical bias in phonological speech errors.

The finding that the stage of phonological encoding is in some way lexically informed has been taken to mean different things within different theoretical frameworks. In fact, the existence of such an effect has been considered an especially powerful piece of evidence for the interactive view of speech production – "more so than any other error phenomenon" (Dell, 1986, p. 300). Within interactive models (e.g., Dell, 1986) it is assumed that the lexical bias effect arises as a result of a bidirectional flow of activation between encoding on a phonological level and on a lexical level. The mechanism proposed is this: When an item is selected on the lexical level, it passes on activation to relevant phonemes on the level of phonological encoding. For example, when the lemma 'cat' is selected, it passes activation to the phoneme /k/. Because of the bidirectional flow of activation in Dell's model, some of this activation that phonemes receive is then fed back up to the lexical level, where the corresponding lexical item receives additional activation: When 'cat' is activated', /k/ becomes activated as well, and vice versa, because the two are connected. This creates a feedback loop of activation – from a selected lexical item, to the corresponding phonemes, and then back to the lexical item, and back to the phonemes, and so on. In this process, the activation of the selected lexical items and corresponding phonemes increases, meaning that they become better candidates for selection. This loop can, of course, only occur when there are relevant items represented at both levels. If there is inappropriate selection of phonemes at the level of phonological encoding, creating a sequence of sounds that is not lexical in the language, then (per definition) there is no corresponding item at the lexical level. In this scenario, no feedback loop can be created. This is why, in the interactive view, an inappropriate phoneme that leads to a non-lexical error outcome is relatively less likely to be selected than an inappropriate phoneme that leads to a lexical error outcome.

In a modular view of speech production, on the other hand, such feedback loops of activation should not be possible. In this view, the Phonological Encoder performs its spellout procedures based only on the lexical selections that were made at the previous stage of encoding, and spits out its output when it is done. There is no interaction from a phonological level to a lexical level and back again during phonological encoding. In this framework, the lexical bias is not a feature of encoding itself, but a side effect of monitoring of covert speech (e.g., Baars et al., 1975; Motley et al., 1982). Baars et al. (1975) posit that the monitor utilizes an explicit check for lexical status on the planned output of the production system, accepting only real words and editing non-lexical items. In a somewhat more conservative version of this explanation, one might assume that the monitor checks simply for any inappropriate and unintended output. Non-lexical output is (almost) always inappropriate and unintended, which is why non-lexical errors are identified and edited by the monitor more often than lexical ones. Lexical outcome errors are more likely to instead 'slip under the radar' in covert monitoring, creating an output bias towards lexical outcomes. The same principle applies for overt monitoring – non-lexical errors in overt speech are more easily detected than lexical errors, which leads to more interruptions of overt non-lexical errors than of lexical errors (Nooteboom, 2005).

2.4 The SLIP-paradigm

As has been noted, speech errors are a crucial type of evidence for processes of speech production. However, they can be highly elusive – after all, most of our utterances are produced as intended. If one wishes to study a particular type of speech error, one has to wait a very long time to collect a sufficient number of them in a naturalistic setting. Other challenges with recording errors in spontaneous speech are inherent differences in detectability for different types of errors, which can make certain types of errors overrepresented in error corpora relative to others (Cutler, 1981).

Eliciting speech errors in an experimental setting mitigates these difficulties and makes it possible to collect a large quantity of a specific type of error and target specific aspects of these errors (see e.g., Stemberger, 1992, for a discussion on differences between naturalistic and experimental speech error data). Baars et al. (1975) were among the first to study phonological speech errors experimentally, by devising what has become known as the SLIP-method (Spoonerisms of Laboratory-Induced Predisposition).

The SLIP-method is aimed at eliciting spoonerisms. The method involves presenting participants with a series of word pairs on a screen, or, in the case of Baars et al.'s original study in 1975, on a memory drum (Chu, 2024), with a very brief exposure period of approximately 1 second. One of these word pairs is the target word pair, which is accompanied by a sound cue signaling that the participant should pronounce the word pair out loud. For example, if a participant is shown the target word pair 'barn door' accompanied with the sound cue, the task is to say 'barn door' out loud, and, hopefully, participants will occasionally slip up and produce 'darn bore' instead, thereby spoonerizing the word pair. If this happens, a speech error has successfully been experimentally elicited. As Baars et al. report, this rarely happens on its own, however. Thus, in order to increase the chances of an error, the target words can be preceded by several bias word pairs which are to be read silently by the participant. The bias pairs are designed to prime the error by exhibiting traits of the desired spoonerism. The target pair 'barn door' might be preceded by the word pairs 'dart board', 'duck bill', and 'dig buck', which will prime the intrusive onset pattern (/d/ and /b/) in favor of the target onset pattern (/b/ and /d/), yielding the spoonerism 'darn bore'.

Baars et al. (1975) used this procedure to investigate the presence of a lexical bias in phonological speech errors. For this purpose, they constructed sets of target pairs of two different types: some that, when spoonerized, had lexical outcomes, such as 'barn door' → 'darn bore', and some that had *non*-lexical outcomes, such as 'dart board' → '*bart *doard'. They hypothesized that if there is a lexical bias in phonological errors, then the rate of errors should change as a function of the lexicality of the expected outcome. In other words, there should be more spoonerisms with lexical outcomes than with non-lexical outcomes. Their results supported this hypothesis, showing that there indeed is an output bias in favor of lexical error outcomes.

Since the genesis of the SLIP-method, the procedure has been reused many times to replicate the lexical bias effect. Various modifications have also been introduced to expand on the understanding of the lexical bias and its origin in processes of speech production. On the one hand, the SLIP-method has been used to test the role of monitoring in speech errors and production in general (Baars et al., 1975; Motley et al., 1982; Wagner-Altendorf et al., 2020). On the other, it has also been used to suggest that the lexical bias is a result of feedback of activation within interactive views on speech production processes (e.g. Dell, 1986; Hartsuiker et al., 2005; Costa et al., 2006; Nooteboom & Quené, 2008). In total, the procedure has been performed on a relatively small number of different languages, namely English (Baars et al., 1975; Motley et al. 1982; Dell, 1986), Spanish (Hartsuiker et al., 2006; Costa et al., 2006) and Dutch (Nooteboom, 2005; Nooteboom & Quené, 2008; Nooteboom & Quené, 2017).

2.5 The current study

As we have seen, speech error evidence is crucial for speech production studies. The lexical bias in particular has been used as evidence in discussions on the interactivity in speech production and the role of the monitor. However, this research has been applied to relatively few languages (English, Dutch, and Spanish). Furthermore, results even within languages have not always been conclusive; for instance, Del Viso et al. (1991) failed to find a lexical bias in a corpus of naturalistic Spanish errors, and Garrett (1975) failed to find it in naturalistic English errors.

The study of a lexical bias effect in speech errors would benefit from being applied to more languages: Assuming that general processes of speech production are universal across all (healthy) adult brains, regardless of which language is being produced, it is important to base assumptions on these processes on data that is consistently indicated to be universal and language-independent. The systematic study of the lexical bias in phonological speech errors has not yet been applied to Swedish, neither in naturalistic nor experimental errors, and the degree of lexical bias in Swedish is not yet known. The present study will investigate the lexical bias in Swedish as an initial step in filling this research gap. This will provide more cross-linguistic evidence on the lexical bias effect as well as much-needed insights into Swedish speech errors. Furthermore, much of the research on the lexical bias effect so far has focused on Germanic languages. Providing data from another Germanic language will go a long way to complete the picture of the lexical bias effect within this language group.

Under the assumptions that 1) there is feedback of activation from a phonological level, and/or 2) there is a monitor that employs a check for lexical status, it follows that there must be a lexical bias in Swedish speech errors. This prediction would hold under either of these theoretical assumptions separately, or under both taken together. Given the experimental evidence so far, I expect to find that the current study will reveal a lexical bias in Swedish speech errors. Specifically, I expect that the data will show fewer non-lexical errors than lexical errors. If, however, the current experiment fails to replicate a lexical bias effect, these theoretical assumptions would need to be reevaluated. Furthermore, it follows from the monitoring account (i.e., assumption 2) that non-lexical errors should be more detectible than lexical ones, not only in covert speech but also in overt speech. Subsequently, we should expect to see not only a lexical bias effect in the error output, but also more overt interruptions and/or corrections of non-lexical errors than of lexical errors.

The present study will experimentally elicit Swedish phonological speech errors in order to investigate their degree of lexical bias. This is done using a modified version of the SLIP-paradigm.

3. Method

3.1 Task

The task in the current study follows the general design of the SLIP-paradigm (see section 2.4). Participants silently read a series of word pairs presented on a computer screen. Some of these word pairs are cued to be spoken aloud, with the hope that participants will occasionally slip up and make a speech error. In the current experiment, each word pair was presented in the center of the screen for a duration of 900 ms, with an inter-stimulus interval (ISI) of 100 ms (following Baars et al., 1975). Alternative durations, with a stimulus exposure time of 700 ms and an ISI of 200 ms, following Costa et al. (2006) and Hartsuiker et al. (2005), were tested during piloting, but proved less effective at eliciting the desired type of errors in comparison to a duration of 900 ms. Upon the conclusion of each trial, a sound cue sounded for 250 ms, signaling for the participant to speak the target word pair out loud. No word pair was visible on screen while the sound cue played, so that participants' responses would not be influenced by any written form. Instead, a '+' sign was displayed, indicating the center of the screen. Participants were instructed to speak their response as soon as they heard the sound cue, but the response period was self-paced. This means that while participants were asked to respond immediately, they could take as long as they wanted to speak and the experiment would not proceed to the next trial until participants gave some keyboard input on the computer. After this, the presentation of the next trial would begin, and so on. The experiment was constructed and conducted using PsychoPy (Peirce et al., 2019).

3.2 Materials

The experimental stimuli are word pairs, which were constructed by me for the present study. Each word pair consists of two monosyllabic Swedish words, for example *moln sylt* ('cloud jam'). There are three types of word pairs to serve different purposes in the experiment: target pairs, bias pairs, and filler pairs. Target pairs are the pairs intended to be spoken aloud and hopefully spoonerized by participants. Bias pairs are intended to prime participants to spoonerize the target pairs. Filler pairs are pairs that serve no purpose except to fill up space between target and bias pairs, creating variation in the trials and making the pattern of bias pairs less obvious. To create these different types of pairs, I compiled lists of monosyllabic Swedish words with different consonantal onsets. A requirement for these words was that onsets were all simple (i.e., one consonant, for example *moln* ‘cloud’), while codas could be either simple (for example *mått* ‘measurement’), complex (i.e., more than one consonant, for example *moln*), or sometimes null (i.e., no consonant, for example *bi* ‘bee’). This was done in order to increase the inventory of possible words. These words were then combined in order to create the three different types of word pairs.

3.2.1 Target pairs

The purpose of the target pairs is to elicit spoonerisms which can reveal a lexical bias. Consequently, the target pairs were constructed in such a way that they can be spoonerized in two different ways: some spoonerize into pairs of lexical items (e.g., ‘could gore’→ ‘good core’) and some spoonerize into pairs of non-lexical items (e.g., ‘cook goes’ → ‘*gook *coes’; examples from Baars et al., 1975). The target pairs were constructed so that for each combination of word onsets, for example /f/ and /s/, there were two very similar pairs, which were variants of each other and differed only in the lexicality of the spoonerism outcome. This means that for the lexical outcome (henceforth LO) target pair *fett synd* ‘fat pity’ (→ *sett fynd* ‘seen find’) there is the non-lexical outcome (henceforth NO) counterpart pair *fäst sytt* ‘fastened sewed’ (→ **säst *fytt*). See Appendix A for the full list of target pairs and their spoonerism outcomes.

There is, to my knowledge, no way to systematically predict which pairings of Swedish words will spoonerize into lexical outcomes. Therefore, creating the target pairs was largely a process of ‘trial and error’, trying various combinations of words until discovering some that met the criteria. Typically, it was more difficult to find word combinations that spoonerized into real words. For this reason, the LO target pairs were created first, and the NO pairs were derived from those by changing the coda. This means that the consonantal onsets as well as the vowels (as far as possible) were kept identical across the two equivalent pairs: If the words in an LO target pair began with the sound sequences *fä-* and *sy-*, then the words in the corresponding NO target pair also began with *fä-* and *sy-*. In some cases, it was difficult to construct LO and NO outcome pairs that were exact equivalents of each other. Therefore, the vowels are sometimes not completely identical between the two target pair variants, for example, *bär tjock* ‘berry fat’ and *bärs tjuv* ‘beer thief’.

The construction of the Swedish target pairs was largely inspired by the target pairs used by Baars et al. (1975, Appendix A). That is, I most often used the same consonantal onsets for the target pairs as they did, in order to give myself a starting point from which to begin pairing the words up and decrease the amount of trial-and-error work. Since Baars et al. used, for example, the target pairs ‘could gore’ (LO) and ‘cook goes’ (NO), which have the consonantal onsets /k/ and /g/, I constructed two Swedish target pairs with the Swedish consonantal onsets /k/ and /g/, and since Baars et al. used the target pairs ‘deep cot’ (LO) and ‘deed cop’ (NO), I constructed two Swedish target pairs with the Swedish consonantal onsets /d/ and /k/, and so on. Like Baars et al., I constructed 80 target pairs in total, of which 40 were LO and the other 40 their NO counterparts. Some target pairs were constructed without deriving the consonantal onsets from the material of Baars et al., specifically those containing the Swedish phonemes /ɕ/, /ɧ/, and /v/.

One important point to note is that the subject of investigation in the current study is centered on the lexical status of the individual words in the target pairs, not the meaningfulness or semantic well-formedness of the words in combination. Some of the two-word sequences I constructed were ungrammatical and meaningless, like *färg bärs* ‘colour beer’, while others were grammatical but meaningless, like *vag dill* ‘vague dill’, while yet others were both grammatical and meaningful, like *fem bär* ‘five berries’. Following Baars et al. (1975), there were no particular restrictions in terms of grammaticality and meaningfulness of the word pairs. This was done in the interest of not restricting the number of possible target pairs, and because the focus of the experiment is the lexical status of individual words, and not the grammaticality or meaningfulness of the two words in combination.

3.2.2 Bias pairs

The purpose of the bias pairs is to interfere phonologically with the target pairs so as to prime spoonerisms. For this reason, they need to exhibit traits of the desired spoonerism outcome of the target pair. This involves displaying a consonantal onset pattern *opposite* to that of a given target pair. To exemplify, for the two target pairs *fett synd* and *fäst sytt*, which have the consonantal onsets /f/ and /s/, bias pairs like *sist fick* ‘last got’, *suck feg* ‘sigh cowardly’ and *sekt fyr* ‘sect lighthouse’ were used, because they display those consonantal onsets in the opposite order and will serve to prime an exchange error.

One criterion for the construction of the bias pair is that no bias pair should contain a word that is identical to any word in any target pair or in any spoonerism outcome. This is to ensure that no spoonerism receives more priming than another, and to ensure that no target pair inadvertently becomes primed in favor of a spoonerism outcome. This would, of course, run counter to the purpose of the experiment. Furthermore, if a bias pair is identical to an expected spoonerism outcome, then it becomes impossible to tell whether or not a participant who produced this word pair actually spoonerized the target pair or is just repeating the bias pair.

3.2.3 Filler pairs

To create the filler pairs, I combined any two words that were not part of any bias pair, target pair, or expected spoonerism outcome. The onset patterns of the filler pairs were designed not to phonologically interfere nor agree with the onsets of the target pairs. For example, the target pair *fett synd* 'fat pity' was preceded by the filler pairs *järn peng* 'iron coin' and *lins minst* 'lens least'.

3.3 Lists and design

The constructed word pairs were embedded into experimental trials and filler trials. The experimental trials each contain one target pair, preceded by a number of bias pairs and filler pairs. Both bias pairs and filler pairs were considered necessary – the bias pairs to increase the likelihood that the target pair is spoonerized, and the filler pairs in order to break up the priming pattern of the bias pairs and make it less obvious to participants. The trials varied in length, consisting of 10, 9, 8, or 7 word pairs in total. The number of bias and filler pairs for each trial varied depending on the length of the trial. The reason for the variation in trial length was to make the appearance of the target pair less predictable, so that participants would have to pay attention to the entire trial.

The experimental trials were constructed under three constraints: Firstly, each trial contains more bias pairs than filler pairs (following the typical procedure of the method, e.g., Hartsuiker et al., 2005, and Nooteboom & Quené, 2008). Secondly, the first bias pair to appear in the trial appears twice, in order to increase the exposure to the phonological interference of the bias pairs and increase the likelihood of a spoonerism (following Baars et al., 1975). Thirdly, the positions immediately adjacent to the target pair are always occupied by a bias pair, and specifically a bias pair that has the same vowel pattern as the target pair. For example, the target pair *fett synd* 'fat pity' was immediately preceded by the bias pair *sekt fyr* 'sect lighthouse', under the assumption that this maximizes interference and increases the number of errors (Baars et al., 1975; Nooteboom & Quené, 2008). According to these constraints, I constructed four different types of experimental trials that were repeated throughout the experiment. These trial types are illustrated in table 1.

Table 1

Experimental trial types

	Experimental trial types				
Position in the trial	Type 1	Type 2	Type 3	Type 4	Example trial type 4
1	Filler	Filler	Filler	Filler	tid jakt
2	Bias	Bias	Bias	Filler	vev jobb
3	Filler	Filler	Filler	Bias	näbb tjut
4	Bias	Filler	Bias	Filler	minst dygn
5	Bias	Bias	Filler	Bias	näbb tjut
6	Target	Bias	Bias	Bias	narr kyss
7	Bias	Target	Bias	Filler	veck sund
8		Bias	Target	Bias	nös köld
9			Bias	Target	köp nöt
10				Bias	nös köld

It is typical in the SLIP-method that the target pair appears in the last position of the trial, and that the sound cue appears at the same time as the presentation of the target pair (see e.g., Baars et al., 1975). However, in the present experiment, the target pair always appears in the *second to last* position. This is because Motley & Baars (1976) found that spoonerism rates increased by 6% when the target pair was not cued immediately, but instead during the presentation of the subsequent pair, presumably due to an increase in task difficulty. Piloting of the present study corroborated an increase in spoonerism rates using this procedure. Thus, this design choice was deemed an appropriate measure to increase the number of errors per participant. Using this procedure, one must control for interference on the target pair from the pair that comes after, considering that this pair will be most recent in participants' memory. For this reason, I chose to always fill the position after the target pair with a bias pair. This final bias pair had to always be identical to the bias pair immediately preceding the target pair (see example in Table 1).

The experiment also contained a number of filler trials. Following Nooteboom & Quené (2008), these were either 5, 4, 3, or 2 word pairs in length, and consisted of random sets of filler pairs. Again, the reason for variation in filler trial length is to make it less predictable when the experimental target pairs will appear. The reason for including very short filler trials, which sometimes consist only of 2 word pairs, is to make sure that participants must pay attention to every word pair in each trial, because even the very first pair in the trial can be cued.

The experimental and filler trials were embedded into two different experimental lists, each containing 40 filler trials and 40 experimental trials (i.e., half of the total experimental trials). Half of participants were exposed to list 1, and the other half was exposed to list 2, meaning that each participant was only exposed to half of the experimental target pairs. This is because for every target pair, there is a version with a lexical outcome (LO), for example *fett synd* 'fat pity' → *sett fynd* 'seen find', and one version with a corresponding non-lexical outcome (NO), for example *fäst sytt* 'fastened sewed' → *säst *fytt (see section 3.2.1). In order to assess the effect of outcome lexicality on the rate of errors, this factor must be isolated from other, potentially confounding effects. It would be suboptimal to present one and the same participant with both the LO variant and the NO variant of a target pair, because regardless of how the experimental material is organized, one variant would necessarily have to appear before the other. This might give rise to ordering effects like practice effects (i.e., performance in the task improves over time), fatigue effects (i.e., performance decreases over time), or carry-over effects (i.e., performance in a trial is affected by the stimuli that participants have been exposed to in an earlier trial). Such effects can all impact the rate of errors and confound the effect of outcome lexicality. For example, if an NO target pair appears, say, 20 trials after its LO counterpart, participants might be more likely to make an error on this NO pair because they are fatigued, or less likely to produce an error because they have had some practice in the task. Furthermore, there is a risk that the phonological priming that participants received during the presentation of the first variant will carry over to the presentation of the second, making them more likely to make an error on the second time of seeing the same priming pattern. To avoid these risks, it is more desirable that half of the participants are exposed to the LO variant of a target pair, while the other half will be exposed to the NO variant of that pair. Consequently, I created two matched lists for a between-subject design. All filler trials and experimental trials were identical in composition and position between the two lists. The only difference between the two lists is that each target pair was placed in the opposite list from its counterpart. For example, if the LO target pair *dör kock* 'dies cook' occupies position 15 in List 1, then the corresponding NO outcome target pair *död kors* 'dead cross' occupies position 15 in List 2. This makes it so that outcome lexicality of the target pairs is the only factor that differs between the two lists, and that the lists are otherwise equivalents of each other.

Each list contains 40 experimental target pairs, 20 of which were supposed to be LO, and 20 of which were supposed to be NO. Because of a minor mistake in the construction of the lists, list 1 was assigned 21 LO pairs and 19 NO pairs, while list 2 was assigned 21 NO pairs and 19 LO pairs. This is because position 23 in list 1, which was supposed to contain an LO pair, accidentally was assigned an NO pair, and vice versa for position 23 in list 2. This means that participants were not exposed to an equal number of LO and NO pairs, but the pairs were still distributed across the lists so that every LO and NO pair was exposed to participants 7 times each.

The ordering of the lists was created through the use of one repeating trial block. Based on the four types of experimental trials and four types of filler trials, I constructed a trial block that contains each of them once, yielding 8 trials per block. This block was repeated 10 times, yielding 80 trials in total. The order of trials was manipulated in this way to ensure that there would not be too many long experimental trials in a row without filler trials to break them up. See Appendix A for the complete experimental material.

3.4 Participants

Fourteen native speakers of Swedish participated in the study. They were recruited by word of mouth. Their ages were in the range of 21-68, with a median age of 26. Six of the participants reported their gender as female, seven as male, and one as other. All participants had English as a second language. In addition to English, five participants reported having a third language, two participants reported a fourth language, and one reported a fifth language, to varying degrees of proficiency. These additional language consisted of German, Dutch, French, Italian, or Spanish. All participants reported Swedish as their first and dominant language. Participants received refreshments as compensation for participating.

3.5 Procedure

Before the experimental session began, participants were informed of the general course and procedure of the experiment. The specific purpose of the experiment was not revealed beforehand. They all signed a form indicating that they give their informed consent to participate in the study (see Appendix B). Thereafter, they filled in the pen-and-paper version of the LEAP-Q questionnaire in Swedish (Marian et al., 2007), detailing their proficiency in their languages. Finally, participants were given written instructions regarding the experimental task (see Appendix C). They were invited to ask questions about the task, and instructions were supplemented orally upon request.

The experiment was run on a laptop, which was placed on a desk in front of the participants. Participants first completed five practice trials with me present, to verify that they had understood the task correctly. This test round consisted only of filler pairs that were cued at random intervals. When participants felt ready to begin the experiment, they were asked to put on headphones equipped with a microphone, which were plugged into a mobile phone that took an audio recording of the main experimental session. The purpose of wearing the headphones was to help isolate potentially interfering noise from the environment. For the main experimental session, I was not present in the room so as not to distract participants or inadvertently influence their responses. Each session took about 15 minutes. After the end of the session, participants were asked to answer follow-up questions regarding how they felt about their performance and what they thought the purpose of the experiment was (Appendix D). After that, participants were debriefed orally, and they were invited to ask questions and discuss.

3.6 Transcription and coding

3.6.1 Transcription

All recorded responses were transcribed using standard Swedish orthography. Non-lexical responses were also transcribed in standard orthography as far as possible. In cases where it was not possible to accurately represent these pronunciations orthographically, they were instead transcribed phonologically. Whenever there was doubt as to what sound sequence a participant had produced, I reviewed the sound recording more thoroughly with the help of spectrogram depictions, and/or consulted another native speaker.

3.6.2 Coding

All responses were coded as one of the following categories. See Table 2 for examples for each category.

- *Full exchange*. Responses that counted as full exchanges were complete spoonerisms of the target pair. In order to be scored as a full exchange, responses had to exactly match the expected spoonerism outcome. If the response deviated from the expected spoonerism outcome even by a single segment, it was not scored as a full exchange.

- *Partial exchange*.[3] Responses were coded as partial exchanges if *one* of the words exactly matched a word in the expected spoonerism outcome.
- *Correct pair*. Responses that were exact and complete repetitions of the target pair were coded as a correct pair. If the response deviated from the target pair even by a single phoneme, it was not scored as correct.
- *Incorrect pair*. Responses that were exact and complete repetitions of some pair that was included somewhere in the current trial, but that was *not* the target pair, were coded as an incorrect pair. If the response deviated from this pair even by a single phoneme, it was instead scored as a mixed pair.
- *Mixed pair*. The mixed pair category includes a wide range of different responses that did not belong to any other category. These were responses consisting of words that were taken from two different pairs in the trial, or from pairs elsewhere in the experiment.
- *Failure to respond*. Responses that were completely missing or which were clearly meta-comments on the task rather than true responses were scored as failures to respond.
- *Innovations*. Participants' responses sometimes contained individual words that were not included anywhere in the experimental material, *or* which were included somewhere in the experiment but that the participant had not yet encountered at the time of producing it. These individual words were coded as 'innovations', since they must have been independently supplied by participants themselves. The innovations could be part either of partial exchanges or mixed pair responses.

Table 2

Examples of response types

Category	Example
Full exchange	*bås gal* → *gås bal*
Partial exchange	*bås gal* → *bås bal, or bås gal* → *gås gal*
Correct	*bås gal* → *bås gal*
Incorrect pair	*gård bak* (bias) and *bås gal* (target) yielding *gård bak*
Mixed error	*gård bak* (bias) and *bås gal* (target) yielding *gård gal*
Failure to respond	…, *jag vet inte* ('I dont know'), *ingen aning* ('no idea'), *nej* ('no'), *fel* ('wrong').

[3] Typically, such errors would be referred to as anticipations or perseverations. However, in the interest of keeping the terminology uniform in the present context, they will be referred to as partial exchanges.

4. Results

Each of the 14 participants was exposed to 40 experimental trials, yielding 560 experimental trial presentations in total. 5,9% of these trials received no response (N=33). Two of the participants were together responsible for most of these failures to respond: one failed to respond 32,5% of the time (N=13) and the other 30% of the time (N=12). Discounting these failures to respond, there were a total of 527 responses to experimental trials.

312 of these 527 responses were correct, meaning that they were exact repetitions of the target pair. On average, participants responded correctly to experimental trials 59,2% of the time. There was considerable inter-participant variation in performance, with a standard deviation of 20,3% from the average rate of correct responses. The participant with the highest number of correct answers responded correctly 92,5% of the time, and the participant with the lowest responded correctly 25,6% of the time.

The remaining responses (N=215) were incorrect in some capacity, meaning that they were not exact repetitions of the target pairs. It is among these incorrect responses that we will find the full and partial exchange errors, which are of primary interest. For the average participant, 3,2% of the incorrect responses were full exchanges, and 7,3% were partial exchanges. In total, there were 24 exchanges, 18 of which were partial exchanges and 6 of which were full exchanges. More commonly than exchanges, the incorrect responses were of the type mixed pairs, which were produced 66,1% of the time on average, or incorrect pairs, which were produced 23,6% of the time on average. In total, there were 47 incorrect pair responses and 144 mixed pair responses. Table 3 shows the response rates of each of these response categories.

Table 3

Rate of responses in each incorrect response category

Error category	N Total	% Total	% Average participant
Full exchange	6	2,8%	3,2% (*SD*=4,3%)
Partial exchange	18	8,4%	7,3% (*SD*=8,4%)
Incorrect Pair	47	21,8%	23,6% (*SD*=20,8%)
Mixed Pair	144	67,1%	66,1% (*SD*=21,7%)
Total	215	100%	100%

Table 4 shows all exchange errors made by participants. The exchanged words are indicated in bold. The only exchange error out of these that was overtly interrupted was the error **sind…*, produced by participant 12.

Table 4

All exchange responses

	Participant[4]								
Partial exchanges	1	2	3	4	5	6	10	12	14
Target pair[5]	Responses								
1.7 kil säck	–	–	–	**sil** käft	**sil** käft	–	–	–	–
2.7 kind seg	–	–	–	–	–	–	–	***sind**… kind seg	–
2.15 död kors	–	–	–	–	–	–	–	–	–
1.18 mur sal	–	–	–	**sur** mark	–	–	–	–	–
2.23 lök rand	–	–	–	–	–	–	–	**rök** *lamn	–
2.25 fett synd	–	–	–	–	–	–	sekt**fynd**	–	–
1.31 lem rund	**rem** lugn	–	–	–	–	–	–	–	–
2.33 bog tål	–	–	–	–	–	–	ton**bål**	–	–
1.39 bås gal	–	–	bås **bal**	–	–	–	–	–	–
2.49 bod gom	–	–	–	–	–	–	–	–	**god** dom
1.50 skär nytt	–	–	–	skär **skytt**	–	–	–	–	–
1.58 pöl kil	kör **pil**	–	luft**pil**	**pil**	**köl** pingst	–	–	–	–
1.60 vakt bås	–	–	–	–	–	–	–	–	–
2.60 valk bår	–	–	–	–	–	–	**balk**bår	–	–
2.65 fukt lem	–	–	–	–	–	–	–	**lukt** *rämn	–
1.66 makt pott	–	–	–	–	**pakt** mun	–	–	–	–
Full exchanges									
Target pair	Responses								
1.2 mård val	–	**vård mal**	–	–	–	–	–	–	–
1.39 bås gal	**gås bal**	–	–	**gås bal**	**gås bal**	**gås bal**	–	–	–
1.47 mygg bäst	–	**bygg mest**	–	–	–	–	–	–	–

4.1 The lexical bias

The degree of lexical bias can be interpreted as the rate of errors produced in the lexical outcome (LO) condition as compared to the rate of errors in the non-lexical outcome (NO) condition. The full exchanges occurred in the LO condition 100% of the time. For the average participant, the partial exchanges occurred in the LO condition 95,2% of the time, and in the NO condition 4,8% of the time. The rate of the other types of responses (correct pair, incorrect pair, and mixed

[4] Participants 7, 8, 9, 11 & 13 are not represented in the table, because they made no exchanges.

[5] The numbering of the target pairs indicates the list to which they belong (1 or 2) followed by their position in that list (1-80). For example, 1.7 refers to the seventh position in the first list. See Appendix A for the experimental material and numbering in its entirety.

pair) were not at all correlated with the lexicality of the target pair's expected exchange outcome. These responses occurred at practically equal rates in the two conditions: on average, 47,9% in the LO condition, and 52,1% in the NO condition. The rate of errors under the different conditions is illustrated in Table 5.

Table 5

Rate of response types in the two outcome lexicality conditions

	N			% Total responses		% Average participant	
Response type	Total	LO	NO	LO	NO	LO	NO
Full exchange	6	6	0	100%	0%	100%	0%
Partial exchange	18	17	1	94,4%	5,6%	95,2% (*SD*=12,6%)	4,8% (*SD*=12,6%)
Full + partial exchanges combined	24	23	1	95,8%	4,2%	96,3% (*SD*=11,1%)	3,7% (*SD*=11,1%)
All other types combined	536	257[6]	279	47,4%	52,6%	47,9% (*SD*=2,8%)	52,1% (*SD*=2,8%)

In addition to the above measure of lexicality, the Innovations in participants' responses proved interesting because they could be either lexical or non-lexical. For example, participant 12 produced the sequences **lamn* and **rämn* which are not exchanges but non-lexical innovations (see Table 4). There were a total of 177 innovations. For the average participant, 85,7% of innovations were lexical, and 14,3% were non-lexical. Table 6 shows the count and proportions of the lexicality of the innovations.

Table 6

Lexicality of the innovations

N			% Average participant	
Lexical	Non-lexical	Total	Lexical	Non-lexical
145	32	177	85,7 % (*SD*=12,8%)	14,3 % (*SD*=12,8%)

[6] There were slightly fewer of these responses in the LO outcome condition as compared to the NO outcome condition. The difference (N=25) almost exactly equals the number of exchanges that occurred in the LO condition, which most likely accounts for the discrepancy.

4.2 Post hoc analyses

A trend that quickly became evident upon reviewing the data is that there appear to be additional factors, other than the lexicality of the outcome, governing the rate of errors. I observed that participants were more likely to spoonerize certain lexical target pairs than other lexical pairs. For example, target pair 1.39 gås bal was fully exchanged by four participants and partially exchanged by one (see Table 4), while the vast majority of experimental target pairs were not exchanged even once. In the post hoc analyses, I tested whether there were some common traits causing certain pairs to be exchanged more often.

After qualitative analysis, I hypothesized that phonetic similarity plays a role. I observed that the onsets of the two words within the target pairs that were subject to exchange (such as *bås* 'booth' and *gal* 'crows' in *bås gal*) were, on average, more similar to each other than those in the pairs that were never subject to exchange (such as *mack* 'gas station' and *tapp* 'tap' in *mack tapp*). To investigate the effect of phonetic similarity on error outcomes, each target pair was retroactively scored according to phonetic similarity between the onsets of the two words in the pairs. Following Bailey & Hahn (2005), each consonant was described according to four features: voicing, place or articulation, manner of articulation, and a sonorant/obstruent distinction. Phonetic similarity was measured in terms of the number of features differing between the onsets. This means that a score of 1 means fewest features differing and highest phonetic similarity, and a score of 4 means most features differing and lowest phonetic similarity. Table 7 shows examples of how onsets were scored.

Table 7

Phonetic similarity scoring: Examples

Target pair	Features word onset 1	Features word onset 2	Features differing
bås gal	/b/ voiced bilabial plosive obstruent	/g/ voiced velar plosive obstruent	1
mack tapp	/m/ voiced bilabial nasal sonorant	/t/ voiceless alveolar plosive obstruent	4
sund rand	/s/ voiceless alveolar fricative obstruent	/r/ voiced alveolar trill sonorant, *or* uvular trill sonorant, *or* uvular fricative obstruent	4, *or* 4, *or* 2

For the rhotic phoneme /r/ in Swedish, there are the multiple allophones. Three realizations were represented among the speakers in the data: [r], [ʀ] and [ʁ]. Therefore, each pair containing the onset /r/ received different scores of phonetic similarity depending on the participant.

This procedure reveals a correlation between phonetic similarity and rate of errors. Specifically, higher degrees of phonetic similarity is correlated with a higher rate of exchanges. The words within the pairs that were subject to exchange, both full and partial, differed from each other by 1,68 features on average. The words involved in partial exchanges differed from each other by 1,82 features on average. The words in the full exchanges were most similar, and differed from each other by 1,33 features on average, and never by more than 2 features. Conversely, the words in the pairs that were not subject to exchange (the correct, incorrect, and mixed pair responses combined) differed from each other by 2,45 features on average. These response categories were fairly uniform in terms of phonetic similarity: The correct pair responses differed by 2,46 features on average, the incorrect responses by 2,52 features, and the mixed pair responses by 2,31 features. Thus, only the full and partial exchanges stand out as being characterized by higher degrees of phonetic similarity. The specific values of phonetic difference for each response category can be seen in Table 8. Values are reported as the average numbers of features differing between onsets per response category, per the average participant.

Table 8

Average number of features differing between word onsets

Response category	Average features differing, per participant
Full exchange	1,33 (*SD*=0,52)
Partial exchange	1,82 (*SD*=0,67)
Correct pair	2,46 (*SD*=0,17)
Incorrect pair	2,52 (*SD*=0,88)
Mixed pair	2,31 (*SD*=0,38)
Full + partial exchanges combined	1,68 (*SD*=0,59)
All others combined	2,45 (*SD*=0,06)

As mentioned, the error *bås gal* → *gås bal* was often produced by participants, which was presumably facilitated by the fact that the two onsets are phonetically similar. However, there is another reason that might have given rise to the higher exchange rate of *bås gal* in particular - namely, that the exchange outcomes *gås* and *bal* might be meaningful to participants in some way. In Swedish, the compound *gåsbal* ‘goose dinner’ refers to a type of formal event during which goose is served for dinner. It appears that this connection between the two words in this error outcome may have played a role in generating exchanges of this target pair specifically. Because of such errors where the semantic relationship between the words within the pairs seemed to affect error rate, I hypothesized that there would also be an effect of grammaticality and/or meaningfulness of the word pairs on the rate of errors. Grammaticality refers to syntactic and morphological well-formedness, that is, whether or not the two words in the pair appear in a permissible order and with the appropriate inflections. Meaningfulness, in the present context, is a measure of the meaningfulness of the *word pairs* and not of individual words. Lexicality is a measure of the lexical status of individual words, that is, whether or not a given sound sequence is recognized as a real word in the language, and all individual lexical items are per definition meaningful. In contrast, a word pair’s meaningfulness here refers to whether or not the two words in combination constitute a meaningful or semantically well-formed sequence. Some word pairs are meaningful sequences, such as *fem bär* ‘five berries’, while others are not, such as *mur sal* ‘wall room’.

In order to investigate the effect of these two factors, three native speakers of Swedish (including myself) scored each target pair and each exchange response according to grammaticality and meaningfulness. These were assigned binary values: 1 for yes, 0 for no. Scorers agreed that no ungrammatical pair was meaningful. Disagreements in scoring were resolved by discussion. Table 9 shows examples of how pairs were scored.

Table 9

Grammaticality and meaningfulness scoring: Examples

Word pair	Grammatical	Meaningful
skär nytt ‘cut new’ (target pair)	1	1
skär skytt ‘pink shot’ (partial exchange)	1	1
bod gom ‘shed palate’ (target pair)	0	0
god dom ‘good verdict’ (partial exchange)	1	1
bås gal ‘booth crows’ (target pair)	0	0
gås bal ‘goose ball’ (full exchange)	0	0
bar koll ‘bare look’ (target pair)	1	0

All noun-noun sequences (e.g., *bod gom* ‘shed palate’ in Table 9), which in English would be grammatical compounds, were not deemed to be grammatical in the Swedish context. This is because in Swedish orthography, compounds are written as one single word not separated by a space. Because all two-word sequences in the material are separated by a space, they are not represented as compounds and therefore not grammatical. This contrast is audible in speech through a difference in prosodic structure. Compounding in Swedish is signaled by a change in stress as well as in word accents, of which Swedish has two, referred to as accent 1 and accent 2. In a pair like *bod gom*, the two individual monosyllabic words are each produced with accent 1. When compounding two monosyllabic words, each consisting of one stressed syllable with accent 1, the prosodic structure changes so that the first syllable of the compound receives primary stress with accent 2 (Bruce, 1977). That participants, like the scorers, did not perceive the noun-noun sequences as being compounds is evidenced by the fact that they (most often) produced all two-word sequences as two separately stressed syllables with accent 1, not a compounded form with accent 2. In the event that a participant did produce an exchange with accent 2, thus signaling that they perceived it to be a compound, the sequence was scored as grammatical.

The effect of these two factors on error outcomes was not as clear as the effects of lexicality and phonetic similarity. There were few grammatical and meaningful target pairs in the experimental material overall, and also few grammatical and meaningful exchange outcomes. This means that for the majority of exchanges, pairs that were both ungrammatical and meaningless turned into error outcomes that were also both ungrammatical and meaningless (e.g., *pöl kil* → *köl pingst*). Participants did this 52,6% of the time on average. The remaining exchanges had outcomes that were either grammatical or both grammatical and meaningful. There were some instances where grammatical and meaningful target pairs exchanged into outcomes that were also grammatical and meaningful, which occurred 5,9% of the time on average. Otherwise, ungrammatical and meaningless target pairs exchanged into grammatical but meaningless outcomes, which occurred 14,8% of the time on average, or into both grammatical and meaningful outcomes, which occurred 26,7% of the time on average. The most interesting finding is that on no occasion did a target pair that was grammatical and meaningful exchange into an ungrammatical and meaningless one. In other words, in 0% of the time was a grammatical and meaningful word pair exchanged if the outcome was ungrammatical and meaningless.

Table 10

Grammaticality and meaningfulness of exchange errors

	Total	Average participant %
Ungrammatical and meaningless target pair to ungrammatical and meaningless exchange	14	52,6%
Ungrammatical and meaningless target pair to grammatical and meaningless exchange	4	14,8%
Ungrammatical and meaningless target pair to grammatical and meaningful exchange	4	26,7%
Grammatical and meaningful target pair to grammatical and meaningful exchange	2	5,9%

5. Discussion

The aim of this thesis was to examine whether there is a lexical bias effect in experimentally elicited Swedish phonological speech errors. The results suggest that there is a clear lexical bias: Out of the total 24 exchanges that occurred in the data, 23 occurred in the lexical outcome condition, showing that, everything else being equal, lexicality of the outcome of a potential error is a very good indicator as to whether or not the error will occur. The presence of the lexical bias effect in Swedish speech errors further indicates the universality of this effect. This confirms what is predicted by both the activation-feedback account and the monitoring account. Furthermore, the only non-lexical error that was uttered was soon interrupted and corrected by the speaker: *kind seg* 'cheek tough'→ *sind... kind seg*. None of the lexical errors were interrupted and corrected in this way. This is indeed what would be expected assuming a monitor that is especially adept at detecting non-lexical error output.

There does appear to be slight cross-linguistic variation in the degree of experimentally elicited lexical bias between the different languages tested so far. For English, Baars et al. (1975) report that 74% of all exchange errors were lexical, and Dell (1986) reports that 62% of all exchange errors were lexical. For Dutch, Nooteboom (2005) reports 66% lexical outcomes, and for Spanish, Hartsuiker et al. (2005) reports 91% lexical outcomes. Out of these, the current Swedish data has the highest degree of lexical bias, with 95,8% of all total exchanges, and 96,3% of all exchanges for the average participant, occurring in the lexical outcome condition. It is doubtful whether this should be attributed to inherently cross-linguistic differences in the lexical bias. Such an interpretation would, following from the two accounts above, imply that the speech production system of a Swedish speaker features more feedback of activation than that of a Dutch or English speaker, or that a Swedish speaker's monitor is more sensitive to non-lexical output. More likely, this difference is a result of slight methodological differences in experimental design and response coding. For example, some studies like Baars et al. (1975) employed a broader definition of exchange errors, and most previous studies using the SLIP-method had a greater number of participants and responses to experimental trials than the present study did. As a result, these studies also elicited more errors overall, and it is very possible that if more errors had been elicited in the current experiment, the proportion of errors in the non-lexical outcome condition would be higher. Further Swedish experiments with more participants could be performed to test this hypothesis.

Interestingly, while the exchange errors exhibited a strong lexical bias in line with the predictions, the form of the exchanges did not follow from what would be expected. In the present study, I elicited more than twice as many partial exchanges (N=18) as full exchanges (N=6). This is similar to Hartsuiker et al. (2005), who elicited 7 full exchanges and 15 partial exchanges. However, based on what would be expected from Levelt's (1989) and Shattuck-Hufnagel's (1979) slots-and-fillers account of speech errors, full exchanges should be *more* common than partial exchanges. This is because partial exchanges require more failures in selection and checkoff than full exchanges do: Full exchanges require the misselection of the first onset phoneme, after which the checkoff mechanism will make sure that the second onset slot will be filled by the remaining onset phoneme (see figure 3 in section 2.2.3) Partial exchanges, on the other hand, require both the misselection of an onset *and* failure of the checkoff mechanism, so that the same onset phoneme can be misselected and appear twice as in *bås gal → bås bal* (see figure 4 in section 2.2.3). This discrepancy between theory and results warrants further consideration of the mechanisms underlying speech errors. Furthermore, most partial exchanges elicited in the current study do not constitute exact anticipation or perseveration errors. Instead, one of the words in the pair was typically exchanged as expected, while the other is erroneous in some other way, for example *pöl kil → köl pingst* or *bod gom → god dom*. In the first example, the second word *pingst* is taken from a bias pair in the trial, and in the second example, the onset of *dom* in the outcome is (presumably) taken from the coda of *bod* in the target pair. Such errors, that is, those that contain some lexical or phonological intrusion seemingly separate from the onset exchange, are plentiful in the data in the form of the mixed errors as well as the partial exchanges. The slots-and-filler account of partial exchanges does not predict that they would so often coincide with other types of errors. One simple explanation for the occurrence of this type of error, that is, all the 'impure' partial exchanges in the data, is that the experimental setting contains many intrusive elements that can work upon the participant at once. The experimental task is demanding, which increases the load on the speech production system and creates more types of errors in one utterance. A very common type of partial exchange is one where the first word in the response is an exchanged form of the target word, while the second word is taken directly from the closest bias pair, for example *kil säck* → *sil käft*, *mur sal* → *sur mark*, *lem rund* → *rem lugn*, and *pöl kil* → *köl pingst*. It is possible that the 'cognitive overload' of the speech production system, which causes the exchange error in the first word, also leads the overburdened Formulator to fail to retrieve and spell out the appropriate second target word and instead replaces it with the one that is most available – that is, the final word in the final bias pair. It is possible that if each target pair had not been immediately followed by a bias pair, as is the more common procedure in the SLIP-paradigm, then there would be fewer of these errors containing lexical intrusions from the bias pairs.

Besides the errors that were actively primed and controlled for, there were a great number of words which we have called innovations. These were words that did not come from the experimental material but were supplied spontaneously by participants, and they could be both lexical and non-lexical. While some of them are likely true speech errors, many of the innovations are likely the result of participants misreading or misremembering the experimental material. Therefore, these innovations must *not* be treated as a corpus of speech errors, and are consequently not sufficient per se to address the question of a lexical bias in Swedish phonological speech errors. However, they still provide some measure of the likelihood of non-lexical utterances in an otherwise lexical context. On average, 84,3% of these innovations were lexical, and 15,7% were non-lexical. That is, when speakers came up with their own responses, these responses were lexical 84,3% of the time. Interestingly, these display a weaker lexical bias than the primed-for exchanges do, which is also in line with previous findings (e.g., Hartsuiker et al., 2005). It is not clear why this is the case. One possibility is that the definitions of non-lexical innovations and lexical innovations were not equal in scope, leading the non-lexical innovations to be overrepresented in the count. In this study, an innovation was defined as a word that did not appear anywhere in the experimental material, or that a participant had not yet encountered in the experiment at the time of producing the word. This means that if a participant produced a word that they had seen, say, as part of a filler pair 50 trials prior, then that word was not considered an innovation but a repetition of experimental material. It is possible that participants produced many lexical items which, for this reason, were not counted as innovations even though they were actually speech errors. On the other hand, no non-lexical innovation was ever part of the experimental material, so these were always easily classified as innovations. Since the definition for the lexical innovations is relatively more narrow, it is possible that fewer lexical innovations were able to be counted and the lexical bias is weaker than it otherwise would have been. A second point is, of course, that the innovations cannot be considered a collection of speech errors. While some of them might be genuine errors, many of them are wholly phonologically unrelated to the target word, which makes them unlikely to be phonological speech errors. Instead, these are more likely a mixed array of misremembered or intentionally misspoken words. There is not necessarily any reason to assume other types of performance mistakes that are not speech errors will display the same degree of lexical bias (if any) as genuine speech errors.

Alongside the lexical bias, post hoc analyses showed that there was also a phoneme similarity effect in the data. This effect seemed to make more similar words more likely to exchange with each other than dissimilar ones. Considering that phoneme similarity was not a controlled factor in the present experiment, the data does not clearly reveal how it interacts with lexicality. Is one of them dominant, meaning that it has a greater effect on the rate of errors than the other? In the first place, we may note that the only error to occur in the non-lexical condition, which was the partial exchange *kind seg* ‘cheek tough’ → *sind… kind seg*, occurred between two onsets that were highly phonetically similar[7] (differing only in terms of place of articulation, and this difference is very slight). This indicates that even if the error outcome is non-lexical, an error may still occur if the segments concerned are sufficiently phonetically similar. However, it seems that the phoneme similarity effect is only minimally capable of overriding the lexical bias: For example, the target pair *bås gal* ‘booth crows’ was subject to some form of exchange 5 times in total, presumably due at least in part to the phonetic similarity between the two onsets. However, the exchanges of the pair *bås gal* occurred only in the lexical outcome condition, and never in the non-lexical condition. This indicates that if the error outcome is non-lexical, the error is still highly unlikely to occur, regardless of phonetic similarity. The lexical bias effect and the phoneme similarity effect appear to operate concurrently with each other.

[7] Swedish ‹k› is typically pronounced not as [k] but approximately as [ɕ] before front vowels, as in *kind* [ɕɪnd].

The lexical bias effect in the data shows that there is some mechanism that either makes lexical errors more likely or makes non-lexical errors less likely. As mentioned, it has been proposed that the lexical bias is the result of either feedback of activation or monitoring. While the current experimental procedure revealed the presence of a lexical bias in Swedish errors, it does not offer many opportunities to tell the underlying causes of the bias apart. The presence of the lexical bias alone can be used to lend support to both of these accounts, but it does not support either account exclusively. Regardless, I will relate how the different pieces of evidence in the present data can be accounted for within both frameworks.

In the first place, it is possible to account for the bias towards lexical outcomes with a monitoring account. There is very little doubt that successful monitoring and editing of errors occurs in covert speech, that is, before the errors are overtly produced. What is less obvious is whether it is truly the case that the monitor is especially adept at identifying non-lexical outcome errors relative to lexical ones. The off-line measures used in the current study do not allow us to witness successful monitoring and editing of covert speech, and it is not possible to tell whether the reason that non-lexical errors are less common is because they are more often detected before articulation. However, monitoring of speech is clearly visible in the form of erroneous utterances that were *overtly* produced and then interrupted and/or repaired. Out of the exchanges in the current data, only one error was interrupted and repaired. This error also happens to be the only exchange to occur in the non-lexical outcome condition: *kind seg* → *sind… kind seg*. Conversely, none of the LO errors were interrupted or repaired. This is probably not mere coincidence: It is probable that it was the non-lexical status of the erroneous utterance that made it particularly easy for the monitor to detect as inappropriate output. It appears that the overtly produced errors that resulted in lexical outcomes were not as easy to detect, judging by the fact that these were not interrupted and/or repaired. It seems, then, that at least overt monitoring favors the detection of non-lexical errors, which is in line with the monitoring account of the lexical bias (Baars et al., 1975; Nooteboom, 2005).

The phoneme similarity effect is also possible to account for with a monitoring account. The monitor can presumably detect non-lexical errors because they are almost certainly inappropriate, in the sense that nonsense utterances are not communicatively helpful and unlikely to be intended. Lexical errors, on the other hand, are less likely to be detected by the monitor because they bear more resemblance to appropriate speech and are more likely to be intended. Similarly, errors that are phonetically similar to the target utterance might be less likely to be detected for that same reason. For example, the error **sind* is extremely similar to the target utterance *kind* 'cheek', which can make the mistake more difficult for the monitor to detect before it is produced. Conversely, *lopp* 'race' and *kok* 'cooking' have onsets that are highly dissimilar, and an exchange between the two might be more detectable for the monitor.

A final trend that was found in the data is that grammatical and meaningful target pairs were never subject to exchange if the error outcome was ungrammatical and meaningless. Conversely, ungrammatical and meaningless target pairs tended to exchange when the outcome was either grammatical or both grammatical and meaningful. This suggest a preference towards grammaticality and meaningfulness in experimentally elicited spoonerisms. This trend is easily explained with the same monitoring mechanism proposed for the lexical bias and phonetic similarity: Sequences that are ungrammatical and meaningless are more easily detected by the monitor than grammatical and meaningful errors are, because these are less likely to be intended by the speaker, leading to an output bias in favor of these traits. 52,6% of the time on average, however, ungrammatical and meaningless target pairs exchanged into outcomes that were also ungrammatical and meaningless. The reason why these target pairs often exchanged even when the outcome was ungrammatical and meaningless can be explained as resulting from a monitor that adapts its standards depending on the intended output: The task of the monitor is to check for output that deviates from the speaker's intention. When the target pair in the task is ungrammatical and meaningless, participants are prepared to intentionally produce ill-formed output. In this situation, it is not useful for the monitor to check for syntactic or semantic well-formedness, because the intention is to produce something ill-formed. Thus, the ungrammatical and meaningless error outcomes can 'slip under the radar', because they do not differ from the target utterance in that regard. When the target pair is grammatical and meaningful, however, participants' intention is to produce something well-formed. In these instances, the monitor is more likely to detect an error outcome that is not well-formed, because it more obviously violates the speaker's intention.

Considering the lexical bias effect, the phonetic similarity effect, and the tendency towards grammaticality and meaningfulness, the monitor can be conceived of as employing a general well-formedness check, rather than a specific check for lexical status as has previously been proposed (e.g., Baars et al., 1975). This would explain why many kinds of anomalous or ill-formed error outcomes, ranging from phonologically illicit, non-lexical, ungrammatical, meaningless, and taboo (which was not investigated in the present study, but see for example Motley et al., 1982), have been found to be relatively less common than neutral or well-formed error outcomes.

The feedback account, on the other hand, posits that the lexical bias effect is a by-product of a spreading activation mechanism. In this view, the lexical bias effect is considered to reflect activation passing both forwards and backwards between a level of lexical selection and phonological encoding, boosting sequences of phonemes that are represented on both levels in favor of sequences that are not represented on the lexical level. In fact, it positively follows from these assumptions that there *must* be a lexical bias in phonological speech errors. In this capacity, the presence of a lexical bias effect in the current data lends support to his view. Similarly, the phoneme similarity effect is also possible to account for in this framework because it is considered to reflect spreading activation, but within the level of phonological encoding (Dell, 1986). In this view, phonemes that have more features in common will pass more activation to each other, making it more likely for an intrusive filler that is similar to the target filler to be selected than an intrusive filler that is not. However, there is no reason why such spreading of activation at one and the same level of encoding is not possible within a discrete framework. Therefore, this piece of evidence has only limited force to support an exclusively interactive view of speech production. Finally, the results suggest that grammatical and meaningful target word pairs are not likely to be exchanged if the outcome is ungrammatical and meaningless. This finding does not lend itself as clearly to the feedback account of output biases as the lexical bias does. It is unclear to me how a mechanism of spreading activation would be sensitive to, for example, the semantic relationship between the words in the word pairs. One possible explanation is that grammatical and meaningful sequences are favored simply because they are more frequent in production, and that the units in these sequences are therefore more strongly connected to each other, leading to more activation spreading between them.

The absence of evidence that speaks exclusively to one or the other accounts of the lexical bias might be interpreted to speak to the likelihood of a combination of the two accounts, as proposed by Hartsuiker et al. (2005) and Nooteboom & Quené (2008).

6. Future research

Considering the results of the current study and previous research, the lexical bias effect appears to hold well cross-linguistically in experimentally elicited errors within the Indo-European language family. An interesting next step would be to investigate the lexical bias in a language unrelated to the languages tested so far. Future studies on the lexical bias effect could also be aimed at investigating more naturally occurring errors in addition to experimental errors. A continuation to the work done in the present thesis could be to investigate naturally occurring speech errors in Swedish. Some work remains to be done not just in regard to the lexical bias, but patterns in Swedish speech errors in general. This would entail a lengthier commitment to collecting a corpus of Swedish speech errors to analyze. Despite the challenging nature of gathering data, such a corpus would have analytic value not just for speech error research but for speech production in general.

A further direction to expand on current results could involve investigating the degree of lexical bias in yet more experimental contexts, in order to shed more light on the psychological origin of the effect. It would be interesting to further investigate and attempt to delimit the role of monitoring more generally in phonological speech errors. In line with previous research in this vein (e.g., Motley et al., 1982; Wagner-Altendorf et al., 2020), this could entail creating more experimental conditions targeted at different types of inappropriate error output, for example taboo word outcomes, in order to reveal the extent of monitoring behind Swedish speech error outcomes.

Finally, it was indicated in the present experiment that factors like phoneme similarity, grammaticality, and meaningfulness played a role in the rate of errors as well. Future experiments would benefit from controlling for these factors, to get a clearer picture of how they interact with the lexical bias effect. In addition to these factors, aspects like word frequency, which were not considered here but which might also be expected to affect rate of error, could be controlled for as well.

7. Conclusion

The present study has indicated a strong lexical bias effect in experimentally elicited Swedish phonological speech errors. The study thus confirms the findings of previous experimental approaches to the lexical bias effect, and is in line with what would be expected based on theoretical assumptions of speech production processes. Additionally, the effect proved relatively stronger than what has been observed so far in studies in other languages. It is not possible to rule out that there are inherent cross-linguistic differences in the degree of a lexical bias effect in phonological speech errors, although this would be unexpected. The results of the current study also revealed that phonetic similarity between the target pairs onsets, as well as grammaticality and meaningfulness of the word pairs, play a role in the rate of errors. There have been two main accounts offered on the origin of the lexical bias effect, namely monitoring and feedback of activation. The findings of the present study can be interpreted in a way that is consistent with both of these accounts, and does not clearly speak to the likelihood of one over the other.

References

Baars, B. J., Motley, M. T., & Mackay, D. G. (1975). Output Editing for Lexical Status in Artificially Elicited Slips of the Tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*, 382–391.

Bailey, T. ., & Hahn, U. (2005). Phoneme Similarity and Confusability. *Journal of Memory and Language*, *52*(3), 339–362.

Boomer, D. S. & Laver, J. (1973). Slips of the tongue. In V.A. Fromkin (Ed.), *Speech errors as linguistic evidence* (p. 120–131). Mouton.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Gleerup.

Chu, Z. (2024). Memory Drum. In *The ECPH Encyclopedia of Psychology*. Springer.

Costa, A., Roelstraete, B., & Hartsuiker, R. J. (2006). The Lexical Bias Effect in Bilingual Speech Production: Evidence for Feedback between Lexical and Sublexical Levels across Languages. *Psychonomic Bulletin & Review: A Journal of the Psychonomic Society, Inc*, *13*(6), 972–977.

Cutler, A. (1982). The reliability of speech error data. In A. Cutler (Ed.), *Slips of the Tongue and Language Production* (p. 7–28). Mouton.

Dell, G. S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, *93*(3), 283–321.

Dell, G. S., & Reich, P. A. (1981). Stages in Sentence Production: An Analysis of Speech Error Data. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 611–629.

Del Viso, S., Igoa, J. M., & García-Albea, J. E. (1991). On the Autonomy of Phonological Encoding: Evidence from Slips of the Tongue in Spanish. *Journal of Psycholinguistic Research*, *20*(3), 161–185.

Fromkin, V. A. (1973). The non-anomalous nature of anomalous utterances. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (p. 215–242). Mouton.

Garrett, M. F. (1975). The analysis of sentence production. In G. Bower (Ed), *The Psychology of Learning and Motivation*. (Vol. 9). Academic Press.

Harley, T. A. (2014). *The psychology of language: from data to theory* (4 ed.). Psychology Press.

Hartsuiker, R. J., Corley, M., & Martensen, H. (2005). The lexical bias effect is modulated by context, but the standard monitoring account doesn't fly: Related beply to Baars et al. (1975). *Journal of Memory and Language*, *52*(1), 58–70.

Hartsuiker, R. J., Anton-Mendez, I., & Roelstraete, B. (2006). Spoonish Spanerisms: A Lexical Bias Effect in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 949–953.

Kerr, E., Ivanova, B., & Strijkers, K. (2023). Lexical access in speech production. In R. J. Hartsuiker & K. Strijkers (Eds.), *Language Production*. Routledge.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. MIT Press.

MacKay, D. G. (1973). Spoonerisms: The structure of errors in the serial order of speech. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (p. 164–194). Mouton.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

Motley, M. T., & Baars, B. J. (1976). Laboratory induction of verbal slips: A new method for psycholinguistic research. *Communication Quarterly*, *24*(2), 28–34.

Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert Formulation and Editing of Anomalies in Speech Production: Evidence from Experimentally Elicited Slips of the Tongue. *Journal of Verbal Learning and Verbal Behavior*, *21*(5), 578–594.

Nooteboom, S. G. (1973). The tongue slips into patterns. In V. A. Fromkin (Ed.), *Speech errors as linguistic evidence* (p. 144–163). Mouton.

Nooteboom, S. G. (2005). Lexical bias revisited: Detecting, rejecting and repairing speech errors in inner speech. *Speech Communication*, *47*(1–2), 43–58.

Nooteboom, S. G., & Quené, H. (2008). Self-Monitoring and Feedback: A New Attempt to Find the Main Cause of Lexical Bias in Phonological Speech Errors. *Journal of Memory and Language*, *58*(3), 837–861.

Nooteboom, S. G., & Quené, H. (2017). Self-monitoring for speech errors: Two-stage detection and repair with and without auditory feedback. *Journal of Memory and Language*, *95*, 19–35.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Poulisse, N. (1999). *Slips of the tongue: speech errors in first and second language production*. J. Benjamins.

Shattuck-Hufnagel, S. (1979). Speech errors as evidence for a serial ordering mechanism in sentence production. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett* (p. 295–342). Erlbaum.

Stemberger, J. P. (1992). The reliability and replicability of naturalistic speech error data: A comparison with experimentally induced errors. In B. J. Baars (Ed.), *Experimental slips and human error: exploring the architecture of volition* (p. 195–215). Plenum Press.

Vigliocco, G., & Hartsuiker, R. J. (2002). The Interplay of Meaning, Sound, and Syntax in Sentence Production. *Psychological Bulletin*, *128*(3), 442–472.

Wagner-Altendorf, T. A., Gottschlich, C., Robert, C., Cirkel, A., Heldmann, M., & Münte, T. F. (2020). The Suppression of Taboo Word Spoonerisms Is Associated With Altered Medial Frontal Negativity: An ERP Study. *Frontiers in Human Neuroscience*, *14*.

Appendix A: Experimental materials

A1: Target pairs

The following list displays all the target pairs used in the current experiment, as well as the expected error outcomes.

Table 1

Target pairs and exchange outcomes

Lexical outcome pairs		Non-lexical outcome pairs	
Target pair	Outcome	Target pair	Outcome[8]
mått sen	sått men	moln sylt	*soln *mylt
mård val	vård mal	mål vagn	*vål *magn
kant lock	lant kock	kalk lån	*lalk *kån
kil säck	sil käck	kind seg	*sind *ɕeg
köp nöt	nöp tjöt	kör nöd	*nör *ɕöd
dov lag	lov dag	dop lav	*lop *dav
räls sked	skälls red	räck sker	*ɧäck *rer
dör kock	kör dock	död kors	*/k/öd *dors
sak tång	tak sång	saft tål	*taft *sål
mur sal	sur mal	mun saft	*sun *maft
sjöss lack	löss schack	skör lapp	*lör *ɧapp
lök rand	rök land	lögn rask	*rögn *lask
fett synd	sett fynd	fäst sytt	*säst *fytt
kant sork	sant kork	kalk son	*salk *kån
sund rand	rund sand	sug rask	*rug *sask
lem rund	rem lund	läpp rytm	*räpp *lytm
bog tål	tog bål	bod tång	*tod *bång
mack tapp	tack mapp	mast tag	*tast *mag
lopp kok	kopp lok	lån koll	*kån *loll
bås gal	gås bal	båt gam	*gåt *bam
bär tjock	kär bock	bärs tjuv	*ɕärs *buv
bit däck	dit bäck	bild del	*dild *bel

[8] The non-lexical error outcomes are transcribed using standard orthography as far as possible. When orthography is not sufficient to convey pronunciation, they are instead loosely phonologically transcribed.

tält folk	fällt tolk	tenn fot	*fenn *tot
mygg bäst	bygg mest	mynt biff	*bynt *miff
bod gom	god bom	bok gång	*gok *bång
skär nytt	när skytt	skäl nyck	*näl *ɧyck
färg bärs	berg färs	fem bär	*bem *fär
bar koll	kar boll	bank kod	*kank *båd
känn mur	män tjur	tjänst mun	*mänst *ɕun
pöl kil	köl pil	pös tjej	*ɕös *pej
valk bår	balk vår	vakt bås	*bakt *vås
dikt seg	sikt deg	disk scen	*sisk *d/eː/n
fukt lem	lukt fem	fusk lek	*lusk *fek
makt pott	pakt mått	mast post	*past *måst
hund koll	kund håll	hus kors	*kus *hors
fin sken	skin fen	fisk sked	*ɧisk *fed
dag vill	vag dill	dam vink	*vam *dink
häst bal	bäst hal	hem barn	*bem *harn
tung sjok	sjung tok	tysk sjuk	*ɧysk *tuk
rätt mör	mätt rör	räls möss	*mäls *röss

A2: List 1

The following lists display the exact content of every trial, experimental and filler, numbered according to the order in which they were presented. Filler trials are indicated by (F), experimental trials with non-word outcomes are indicated by (N), and experimental trials with lexical outcomes by (L). Word pairs are separated from each other by commas.

List 1: complete materials

Trial	Word pairs in order of presentation
1 (N)	gadd vakt, påsk gadd, tofs sjal, sekt mark, sekt mark, såg minst, pust natt, sug mur, sond märg, moln sylt, sond märg
2 (L)	växt mor, sol sund, väck nåd, tjat rak, växt mor, vev mål, våt mark, mård val, våt mark
3 (F)	dans namn, föl skylt
4 (N)	luft kung, post vurm, luft kung, fick ryck, jakt hud, loft kaj, lins keps, lamm kår, kalk lån, lamm kår
5 (F)	jakt häck, fall vits, nåd ben, tofs pall
6 (F)	jakt vurm, dill helg, föl nåd
7 (L)	såg kyl, lim holk, såg kyl, natt väv, fall bak, sist käft, kil säck, sist käft
8 (F)	feg hus, tag pöl, pyr häck, fas golv, makt bi
9 (N)	tid jakt, vev jobb, veck sund, näbb tjut, näbb tjut, narr kyss, minst dygn, noll kjol, nös köld, kör nöd, nös köld

10 (L) lyft dolk, pank hyr, tid pott, bit tand, lyft dolk, läck dans, loft dal, dov lag, loft dal

11 (F) pung vev, galt sjok

12 (N) skikt regn, jul tjat, skikt regn, tapp bok, luft rep, skjut rönn, skön rop, skägg rep, räck sker, skägg rep

13 (F) väv hot, pank gul, bi sund, vakt veck

14 (F) beige kyl, tid kyss, den rytm

15 (L) kung där, jakt näst, kung där, falk mark, själ sund, kö dolk, dör kock, kö dolk

16 (F) veck näbb, sund bok, suck helg, gem suck, golv narr

17 (N) hög veck, fas tid, dess kyss, tofs suck, tofs suck, töm sill, påsk väv, ton sekt, tand såg, saft tål, tand såg

18 (L) sol maj, bok kyl, häck jobb, rytm halv, sol maj, sax mån, suck mark, mur sal, suck mark

19 (F) pung hyfs, den dess

20 (N) lyft skydd, natt fall, lyft skydd, våg feg, hus rep, löd skärp, lins schakt, löv sjal, skör lapp, löv sjal

21 (F) påsk jakt, väv vits, bi nog, tand pott

22 (F) vits pall, halv gul, lår pott

23 (N) rang loft, famn mer, rang loft, suck jul, hingst dikt, röd lamm, lögn rask, röd lamm

24 (F) golv falk, jakt dikt, suck pung, hyr helg, djup pust

25 (N) lins minst, järn skäl, tid feg, sank famn, sank famn, sist fick, järn peng, suck feg, sekt fyr, fäst sytt, sekt fyr

26 (L) sond kam, lår bock, pyr tapp, holk mer, sond kam, sed red, sax korp, kant sork, sax korp

27 (F) den bi, sug lår

28 (N) råd suck, föl fast, råd suck, bad pust, tand fall, rang sol, sill rönn, rus sank, sug rask, rus sank

29 (F) natt lins, dans fast, sjok vakt, bad bit

30 (F) vits byst, sug den, jobb pigg

31 (L) ryck lam, bit pyr, ryck lam, lins pust, beige sekt, rest lugn, lem rund, rest lugn

32 (F) makt fas, rytm tenn, tenn napp, kung veck, pyr bok

33 (N) pöl hymn, nöd son, helg pingst, töm bygd, töm bygd, tyg beige, son hot, tid band, ton bård, bod tång, ton bård

34 (L) tåg mord, kyss nog, djup gul, pöl hyr, tåg mord, tofs märg, tand mall, mack tapp, tand mall

35 (F) halv visp, väv sug

36 (N) keps lam, vits bit, keps lam, nåd halv, fall dikt, kår löd, kall lins, korg log, lån koll, korg log

37 (F) hymn noll, skydd sjal, fall pyr, dess son

38 (F) pott dikt, post pott, suck jakt

39 (L) gul beige, dans seg, gul beige, tag hud, jobb kund, gård bak, bås gal, gård bak

40 (F) namn hyfs, pigg vits, haj pall, tenn pust, pust feg

41 (N) sjal jakt, tag pigg, nog suck, kyl beige, kyl beige, kört besk, pank hymn, käft bård, käk bur, bärs tjuv, käk bur

42 (L) dolk bak, pöl noll, tenn pank, hud vurm, dolk bak, dess byst, dill besk, bit däck, dill besk

43 (F) bok peng, gadd vakt

44 (N) föl tand, råd hud, föl tand, hymn galt, kung pust, fick ton, fog tyg, fäst torg, tenn fot, fäst torg

45 (F) hot haj, sjal hus, bad byst, narr gul

46 (F) fas dans, sekt bad, son rytm

47 (L) borr märg, tag näbb, borr märg, galt haj, sjok lins, byst mord, mygg bäst, byst mord

48 (F) hingst makt, feg sjok, galt råd, skikt pall, bad vurm

49 (N) kyss pank, haj föl, hingst pingst, gul bomb, gul bomb, galt besk, fall dans, gem byst, gos borg, bok gång, gos borg

50 (L) napp sjal, rak kung, vurm gul, pigg dikt, napp sjal, nåd skikt, näbb skydd, skär nytt, näbb skydd

51 (F) råd galt, röd vakt

52 (N) bomb föl, galt nog, bomb föl, dikt gem, son pott, besk far, bård feg, bädd fäst, fem bär, bädd fäst

53 (F) natt noll, våg feg, skikt jakt, fick varg

54 (F) kyss vev, byst lår, napp makt

55 (L)	korp besk, hyr tapp, korp besk, rytm pott, häck djup, kamp borg, bar koll, kamp borg
56 (F)	byst hyr, dans helg, hud jobb, bit hud, växt vakt
57 (N)	famn hus, nåd sol, jul rop, mer kyss, mer kyss, mord kjol, famn kyss, mall köld, märr tjut, tjänst mun, märr tjut
58 (L)	käft pingst, dill fast, vits gadd, luft seg, käft pingst, käk pank, kört pyr, pöl kil, kört pyr
59 (F)	son näbb, fas natt
60 (N)	bomb visp, noll pynt, bomb visp, mer hot, tjat hus, beige vits, borr växt, band våt, vakt bås, band våt
61 (F)	golv skydd, tand röd, helg pigg, sist väck
62 (F)	kyl pust, tenn visp, makt tenn
63 (L)	suck dörr, fel djup, suck dörr, rak hög, luft järn, sist det, dikt seg, sist det
64 (F)	väck feg, tag påsk, fick gagn, sist väck, haj pyr
65 (N)	bad gul, hymn rop, seg rop, lamm föl, lamm föl, lins falk, ryck skärp, löd fyr, luft fäst, fusk lek, luft fäst
66 (L)	pust minst, kyss tand, beige galt, rop kung, pust minst, päls mark, pank mobb, makt pott, pank mobb
67 (F)	jakt peng, rop dygn
68 (N)	kål hymn, skydd jobb, kål hymn, lins ben, växt hymn, kaj hingst, korg hud, kung holk, hus kors, kung holk
69 (F)	tofs fast, sekt falk, pung råd, skydd tand
70 (F)	råd dikt, fick dygn, jakt pank
71 (L)	skön famn, pöl tofs, skön famn, föl den, röd djup, skikt fel, fin sken, skikt fel
72 (F)	pöl järn, föl lins, häck veck, sjal tag, pingst hus
73 (N)	hyr golv, pingst gem, järn gul, vurm dolk, vurm dolk, våg där, näbb pank, varg dörr, var din, dam vink, var din
74 (L)	bård harv, gadd vakt, haj makt, påsk sjal, bård harv, beige halv, besk hak, häst bal, besk hak
75 (F)	napp noll, keps växt
76 (N)	skägg töm, kung häck, skägg töm, rak sol, hymn natt, skärp tyg, skön tåg, skjut ton, tull sjuk, skjut ton
77 (F)	nos sund, suck jul, noll dikt, sund pingst
78 (F)	hot namn, pank regn, seg bygd
79 (L)	mobb rop, påsk gadd, mobb rop, pust natt, gem dolk, märr röd, rätt mör, märr röd
80 (F)	kung fall, skylt röd, bok napp, näbb pott, helg pust

A3: List 2

List 2: complete materials

Trial	Word pairs in order of presentation
1 (L)	gadd vakt, påsk gadd, tofs sjal, sekt mark, sekt mark, såg minst, pust natt, sug mur, sond märg, mått sen, sond märg
2 (N)	växt mor, sol sund, väck nåd, tjat rak, växt mor, vev mål, våt mark, mål vagn, våt mark
3 (F)	dans namn, föl skylt
4 (L)	luft kung, post vurm, luft kung, fick ryck, jakt hud, loft kaj, lins keps, lamm kår, kant lock, lamm kår
5 (F)	jakt häck, fall vits, nåd ben, tofs pall
6 (F)	jakt vurm, dill helg, föl nåd
7 (N)	såg kyl, lim holk, såg kyl, natt väv, fall bak, sist käft, kind seg, sist käft
8 (F)	feg hus, tag pöl, pyr häck, fas golv, makt bi
9 (L)	tid jakt, vev jobb, veck sund, näbb tjut, näbb tjut, narr kyss, minst dygn, noll kjol, nös köld, köp nöt, nös köld
10 (N)	lyft dolk, pank hyr, tid pott, bit tand, lyft dolk, läck dans, loft dal, dop lav, loft dal
11 (F)	pung vev, galt sjok
12 (L)	skikt regn, jul tjat, skikt regn, tapp bok, luft rep, skjut rönn, skön rop, skägg rep, räls sked, skägg rep

13 (F)	väv hot, pank gul, bi sund, vakt veck

14 (F)	beige kyl, tid kyss, den rytm

15 (N)	kung där, jakt näst, kung där, falk mark, själ sund, kö dolk, död kors, kö dolk

16 (F)	veck näbb, sund bok, suck helg, gem suck, golv narr

17 (L)	hög veck, fas tid, dess kyss, tofs suck, tofs suck, töm sill, påsk väv, ton sekt, tand såg, sak tång, tand såg

18 (N)	sol maj, bok kyl, häck jobb, rytm halv, sol maj, sax mån, suck mark, mun saft, suck mark

19 (F)	pung hyfs, den dess

20 (L)	lyft skydd, natt fall, lyft skydd, våg feg, hus rep, löd skärp, lins schakt, löv sjal, sjöss lack, löv sjal

21 (F)	påsk jakt, väv vits, bi nog, tand pott

22 (F)	vits pall, halv gul, lår pott

23 (L)	rang loft, famn mer, rang loft, suck jul, hingst dikt, röd lamm, lök rand, röd lamm

24 (F)	golv falk, jakt dikt, suck pung, hyr helg, djup pust

25 (L)	lins minst, järn skäl, tid feg, sank famn, sank famn, sist fick, järn peng, suck feg, sekt fyr, fett synd, sekt fyr

26 (N)	sond kam, lår bock, pyr tapp, holk mer, sond kam, sed red, sax korp, kalk son, sax korp

27 (F)	den bi, sug lår

28 (L)	råd suck, föl fast, råd suck, bad pust, tand fall, rang sol, sill rönn, rus sank, sund rand, rus sank

29 (F)	natt lins, dans fast, sjok vakt, bad bit

30 (F)	vits byst, sug den, jobb pigg

31 (N)	ryck lam, bit pyr, ryck lam, lins pust, beige sekt, rest lugn, läpp rytm, rest lugn

32 (F)	makt fas, rytm tenn, tenn napp, kung veck, pyr bok

33 (L)	pöl hymn, nöd son, helg pingst, töm bygd, töm bygd, tyg beige, son hot, tid band, ton bård, bog tål, ton bård

34 (N)	tåg mord, kyss nog, djup gul, pöl hyr, tåg mord, tofs märg, tand mall, mast tag, tand mall

35 (F)	halv visp, väv sug

36 (L)	keps lam, vits bit, keps lam, nåd halv, fall dikt, kår löd, kall lins, korg log, lopp kok, korg log

37 (F)	hymn noll, skydd sjal, fall pyr, dess son

38 (F)	pott dikt, post pott, suck jakt

39 (N)	gul beige, dans seg, gul beige, tag hud, jobb kund, gård bak, båt gam, gård bak

40 (F)	namn hyfs, pigg vits, haj pall, tenn pust, pust feg

41 (L)	sjal jakt, tag pigg, nog suck, kyl beige, kyl beige, kört besk, pank hymn, käft bård, käk bur, bär tjock, käk bur

42 (N)	dolk bak, pöl noll, tenn pank, hud vurm, dolk bak, dess byst, dill besk, bild del, dill besk

43 (F)	bok peng, gadd vakt

44 (L)	föl tand, råd hud, föl tand, hymn galt, kung pust, fick ton, fog tyg, fäst torg, tält folk, fäst torg

45 (F)	hot haj, sjal hus, bad byst, narr gul

46 (F)	fas dans, sekt bad, son rytm

47 (N)	borr märg, tag näbb, borr märg, galt haj, sjok lins, byst mord, mynt biff, byst mord

48 (F)	hingst makt, feg sjok, galt råd, skikt pall, bad vurm

49 (L)	kyss pank, haj föl, hingst pingst, gul bomb, gul bomb, galt besk, fall dans, gem byst, gos borg, bod gom, gos borg

50 (N)	napp sjal, rak kung, vurm gul, pigg dikt, napp sjal, nåd skikt, näbb skydd, skäl nyck, näbb skydd

51 (F)	råd galt, röd vakt

52 (L)	bomb föl, galt nog, bomb föl, dikt gem, son pott, besk far, bård feg, bädd fäst, färg bärs, bädd fäst

53 (F)	natt noll, våg feg, skikt jakt, fick varg

54 (F)	kyss vev, byst lår, napp makt

55 (N)	korp besk, hyr tapp, korp besk, rytm pott, häck djup, kamp borg, band kod, kamp borg

56 (F)	byst hyr, dans helg, hud jobb, bit hud, växt vakt

57 (L)	famn hus, nåd sol, jul rop, mer kyss, mer kyss, mord kjol, famn kyss, mall köld, märr tjut, känn mur, märr tjut

58 (N)	käft pingst, dill fast, vits gadd, luft seg, käft pingst, käk pank, kört pyr, pös tjej, kört pyr
59 (F)	son näbb, fas natt
60 (L)	bomb visp, noll pynt, bomb visp, mer hot, tjat hus, beige vits, borr växt, band våt, valk bår, band våt
61 (F)	golv skydd, tand röd, helg pigg, sist väck
62 (F)	kyl pust, tenn visp, makt tenn
63 (N)	suck dörr, fel djup, suck dörr, rak hög, luft järn, sist det, disk scen, sist det
64 (F)	väck feg, tag påsk, fick gagn, sist väck, haj pyr
65 (L)	bad gul, hymn rop, seg rop, lamm föl, lamm föl, lins falk, ryck skärp, löd fyr, luft fäst, fukt lem, luft fäst
66 (N)	pust minst, kyss tand, beige galt, rop kung, pust minst, päls mark, pank mobb, mast post, pank mobb
67 (F)	jakt peng, rop dygn
68 (L)	kål hymn, skydd jobb, kål hymn, lins ben, växt hymn, kaj hingst, korg hud, kung holk, hund koll, kung holk
69 (F)	tofs fast, sekt falk, pung råd, skydd tand
70 (F)	råd dikt, fick dygn, jakt pank
71 (N)	skön famn, pöl tofs, skön famn, föl den, röd djup, skikt fel, fisk sked, skikt fel
72 (F)	pöl järn, föl lins, häck veck, sjal tag, pingst hus
73 (L)	hyr golv, pingst gem, järn gul, vurm dolk, vurm dolk, våg där, näbb pank, varg dörr, var din, dag vill, var din
74 (N)	bård harv, gadd vakt, haj makt, påsk sjal, bård harv, beige halv, besk hak, hem barn, besk hak
75 (F)	napp noll, keps växt
76 (L)	skägg töm, kung häck, skägg töm, rak sol, hymn natt, skärp tyg, skön tåg, skjut ton, tung sjok, skjut ton
77 (F)	nos sund, suck jul, noll dikt, sund pingst
78 (F)	hot namn, pank regn, seg bygd
79 (N)	mobb rop, påsk gadd, mobb rop, pust natt, gem dolk, märr röd, räls möss, märr röd
80 (F)	kung fall, skylt röd, bok napp, näbb pott, helg pust

Appendix B: Consent form

Samtyckesblankett: Kandidatuppsats i allmän lingvistik, Hanna Roskvist

---

Härmed intygar jag att jag har fått muntlig och skriftlig information om studien och har fått möjlighet att ställa frågor. Jag samtycker till att ljudupptagningarna inspelade under dagens experimenttillfälle används av Hanna Roskvist till vetenskaplig analys i forskningsändamål.

Jag har fått information om att de uppgifter som samlas in om mig kommer att behandlas konfidentiellt, på ett sådant sätt att min identitet inte kommer att avslöjas för obehöriga. Min identitet kommer endast att vara känd för Hanna Roskvist och förblir totalt anonym för alla andra.

Jag är medveten om att min medverkan är helt frivillig och att jag när som helst och utan närmare förklaring kan avbryta mitt deltagande.

Namn Underskrift Datum

Appendix C: Instructions

Välkommen till experimentet.

Under experimentets gång kommer en följd av ord dyka upp på skärmen. Orden kommer att dyka upp två i taget. Din uppgift är att läsa alla ord *tyst* i huvudet. I slumpmässiga intervaller under experimentet kommer en ljudsignal att spelas upp. När du hör signalen ska du uttala orden som du såg *näst sist*. ***Tala så fort som möjligt***.

Här följer ett exempel:

- katt hund
- fågel fisk
- mus häst
- [LJUDSIGNAL]

Om du ser den här följden av ord ska du alltså säga *fågel fisk* när du hör ljudsignalen.

Innan experimentets huvuddel börjar kommer du få köra en testomgång för att säkerställa att du har förstått instruktionerna.

Appendix D: Follow-up questions

Uppföljande frågor

Hur upplevde du att det gick?

Vad tror du experimentet gick ut på?

Övriga kommentarer, kritik och feedback: