Can spontaneous spoken language disfluencies help describe syntactic dependencies? An empirical study

M. Zakaria KURDI
Department of Computer Science, University of Lynchburg,
Lynchburg, VA
kurdi_m@lynchburg.eduAbstract

Abstract

This paper explores the correlations between key syntactic dependencies and the occurrence of simple spoken language disfluencies such as filled pauses and incomplete words. The working hypothesis here is that interruptions caused by these phenomena are more likely to happen between weakly connected words from a syntactic point of view than between strongly connected ones. The obtained results show significant patterns with the regard to key syntactic phenomena, like confirming the positive correlation between the frequency of disfluencies and multiples measures of syntactic complexity. In addition, they show that there is a stronger relationship between the verb and its subject than with its object, which confirms the idea of a hierarchical incrementality. Also, this work uncovered an interesting role played by a verb particle as a syntactic delimiter of some verb complements. Finally, the interruptions by disfluencies patterns show that verbs have a more privileged relationship with their preposition compared to the object Noun Phrase (NP).

1 Introduction

This paper explores the way speech stream is interrupted by *simple spoken language disfluencies* (from now disfluencies) such as filled pauses and incomplete words (Kurdi, 2016). It aims to shed light on language planning during the language generation process through the window of disfluencies. One of the key questions this work tries to answer is how tightly related are some syntactic components within an utterance. The underlying hypothesis here is that tightly related components are planned together and consequently less likely interrupted by a disfluency.

Another contribution of this work is to provide a numeric value to describe the strength of the linguistic connection between two words, as this study is conducted at the scale of an entire corpus. Please note that the linguistic and cognitive validity of the existing statistical models to describe the strength of a dependency, based on the co-occurrence of words and structures, is highly disputed by many linguists like Chomsky. A basic argument against such models is that a rare structure can be as grammatical as a frequently used one. Hence, the potential applications of this work within the area of NLP can range from syntactic disambiguation to the reranking of speech recognition N best hypotheses.

In previous research, disfluencies were explored from multiple points of views. For example, (Carbonell and Hayes, 1984), (Heeman, 1999), (Core, 1999), and (Kurdi, 2002) investigated this relation within the context of spoken language parsing. In the psycholinguistics literature, several models stressed the role of syntax in the process of language production and planning. For instance, serial processing models of language production such as Fromkin's five stage model (Fromkin, 1973), Garrett's model (Garrett, 1980), (Garrett, 1988), and Bock and Levelt's model (Bock and Levelt, 1994) assume the existence of an explicit module for syntactic processing to which they attribute different names and functional roles. In connectionists models, such as Dells' model (Dell et al, 1999), all knowledge levels interact with each other, with the lexicon playing a central role in this process. When a word is selected all the phonological, morphological and syntactic features related to its constituents are also activated and propagated to the context, contributing to activate new words. This suggests that syntactic dependencies between words play a key role in the process of spoken language production.

Besides, several previous works stipulate that self-monitoring plays a key role in language production. In particular, Levelt's Perceptual Loop Theory (PLT) suggests that there exist two modes of monitoring

Proceedings of the First International Workshop on Language Cognition and Computational Models, pages 75–84
Santa Fe, New Mexico, United States, August 20, 2018.
https://doi.org/10.18653/v1/P17

(Levelt, 1983). The first one consists of monitoring internal unproduced speech which consists of checking one's planned formulation silently. Similar to the process of listening to other's speech, the external monitoring, on the other hand, consists of monitoring one's speech by ears. Both processes, involve treatment by the speech comprehension system, which covers both the semantic and syntactic aspects of language. Some more recent works such as the ones of (Nozari, Dell, & Schwartz, 2011) and (Hartsuiker and Herman, 2001) argue for internal monitoring based on competition between representations within the language production system without the intervention of the comprehension system. It is hard to see how these new studies can contradict the idea of the intervention of syntax within the monitoring system for the following reasons. First, these studies focused on low-level linguistic phenomena such as word production and do not take into consideration the syntactic structure. In addition, self-repair can be motivated to correct syntactic errors. Likewise, many works have indicated that discourse, syntax, and prosody play an important role within language planning (see (Wagner, 2016) for a review of these works).

Furthermore, multiple works have shown that there is a correlation between language complexity in general and production of disfluency (McLaughlin and Cullinan, 1989), (Haynes, Hood, 1978). More specifically, syntactic complexity is linked to frequency of production of disfluencies (Gordon and Luper, 1989), (Logan and Lasalle, 1999) disfluency initiation times (Ferreira, 1991). Besides, (Boomer, 1965) reported that filled pauses tend to appear between the first and the second word of a clause, suggesting that this may be related to the syntactic planning of the utterance. Some other works focused on syntactic planning and disfluency within the context of foreign language (Rose, 2017).

A question one could ask about the generation, the planning or the monitoring processes is the following. Which syntactic unit is used by these processes? Some studies suggested that clause (or simple sentence) plays a key role in this process (Ford and Holmes, 1978), (Rose, 2017), while others stipulated that structures like LTAG trees are used (Ferreira, 2000). In addition, Levelt, in his extension of Dell's three level model, assumes that the grammar encoding is done within the lemma-stratum module where processing is based on syntactic features of individual words such as tense for the verbs (Levelt et al, 1999).

2 Methodology

2.1 Hypotheses

The working assumption in this paper is that the locations of the interruptions of speech flow by disfluencies are related to the syntactic dependencies within the utterance. For example, if the interruption happens rarely within a given context (e.g. between two morphological categories, like a determiner and a noun DT NN) we assume that the components involved in this context are strongly connected and vice versa.

This fundamental assumption leads to the following four hypotheses:

i. Disfluencies are the reflection of a heavy cognitive processing (Lindström, 2008). Hence, it is more likely that disfluencies occur in a more syntactically complex utterance.
ii. Given their shared features, verbs are more tightly connected to their subject than to their object. This means that it is less likely to observe an interruption between a verb and its subject than between a verb and its object.
iii. The relation between particles and verbs is so tight morphologically. From a semantic point of view, a particle may change the meaning of some verbs. In addition, it is hypothesized here that verb particles play a key syntactic role in planning and delimiting some of the verb arguments.
iv. Given the privileged relationship between the verb and its preposition, it is hypothesized that interruptions between the verb and the preposition are less likely than between the preposition and the subsequent Noun Phrase (NP).

2.2 Corpus

The Trains Corpus (Heeman and Allen, 1995) was used because of the quality of transcription and reasonable size: 98 dialogues with 34 different speakers and 5,900 speaker turns. Unlike other spoken language corpora, the task is complex which creates more opportunities for producing disfluencies. After a comparative study with a portion of the switchboard corpus (Meteer, 1995), it was possible to observe that the disfluencies available in the Trains Corpus are similar to the ones in the Switchboard Corpus.

2.3 Data annotation

The disfluencies are annotated using the scheme adopted in (Kurdi, 2003). Given that the focus of this work is about syntax, are adopted the following criteria for defining an interruption of the utterance flow. First, filled pauses and incomplete words such as *hum* and *prob*- are the obvious indicators. Some prosodic events such as silence (unfilled pauses) were not considered. The problem with silence is that it is hard to mark with high accuracy given the individual differences between speakers' pace. Also, speakers may take a short pause for the sake of breathing, a rather physiological event. Finally, silence markers are likely to be accompanied by one or more of the adopted interruption indicators. Are also excluded contextual and physiological events such as breadth and laugher as they are not necessarily related to language planning.

2.4 Interruption rate

To provide a probability-like measure of the connectivity rate, Interruption Rate (IR) is adopted. It is calculated using equation 1 where c(x) means the count of *x*:

$$\text{interruption\_rate(n-gram}_i) = \frac{\text{c(interrupted n-gram}_i)}{\text{c(all occurences of n-gram}_i)} \quad (1)$$

To observe the interruption patterns, two programs are implemented. A statistical part-of-speech (POS) tagger based on a cascade of n-grams trained on the Penn tree bank. To correct the errors with this tagger, is also implemented a post-processing module. It corrects two types of errors: generic and corpus specific errors. For example, is used a rule that would retag all the auxiliary verbs as MD when they are used before a verb. An example of a corpus specific error is the word *Corning* which is only used in the corpus as a proper noun (a city in the state of New York) but the statistical tagger sometimes tags it as a verb. The tag set adopted is inspired by the Penn treebank[1].

The implemented program provides a *raw* interruption rate. Given that n-grams provide only a sequencing of POS tags, which does not necessarily reflect a relation of dependency, all the sequences are checked *manually*. Are considered as *syntactic* interruptions only those that occur between syntactically related words. For example, in the sequence (DT NN VB) such as the one in *okay so just a second **uh** let me see -what time (...)*. The interruption here, by the filled pause *uh*, is not between syntactically dependent words as the sequence *a second* belongs to a different utterance and is not a subject or an object of the verb *see*. Therefore, it is not counted as a syntactic interruption. However, in the sequence (DT NN VB) in *we do not have two trains uh trying to cross (..)* there is a syntactic interruption as *two trains* is the subject of the verb *trying*.

The IR of a specific bigram is compared to the IR of the general bigram (XX), which is .026. The bigram (XX) is made with average IR of all observed sequences of two POS tags in the corpus. Similarly, the IR of a specific trigram is compared to the interruption of the IR of the general trigram pattern (XXX), which is .049.

3 Results

3.1 Disfluencies and utterance syntactic complexity

Several works in the literature have reported that the chance of disfluency production increases with the increase of conceptual or linguistic difficulty of the utterance. In this study, five different measures of syntactic complexity were considered and their correlation with the number of disfluencies within the utterance was calculated (see (Kurdi, 2017) for more information about these measures). The measures involving phrases and the depth of the parsing tree were calculated with the Sandford parser[2].

As seen in Table 1, the syntactic complexity indices and the number of disfluencies have a statistically significant positive correlation, meaning increases in syntactic complexity of an utterance were correlated with increases in the number of disfluencies. The smallest correlation is with the number of

[1] https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html
[2] https://nlp.stanford.edu/software/lex-parser.shtml

VB while the largest is with the length of the utterance. This difference between the two correlations remains limited, as it is about 14% of the total value.

Complexity measure	Pearson's correlation with the number of disfluencies
Number of phrases in the utterance	.286
Depth of parsing tree of the utterance	.249
Mean length of the phrases	.276
Number of verbal phrases	.247
Length of the utterance	.289

Table 1 Correlations between the number of disfluencies per utterance and five indices of syntactic complexity, for all the correlations [N=5020, p<.0001]

3.2 Connectivity between the verb and its subject and object

In English, where the canonical order is Subject-Verb-Object (SVO), the verb is the heart of the sentence. The relation between the verb and its subject is privileged because of their shared syntactic features, as they agree in number. The question now is the following. What is the effect of this privileged relationship on the strength of the dependency between the verb and its subject? To answer this question, a two-fold process was carried out. First, the left and right connectivity of the verb is calculated through the patterns (X VB) and (VB X). The first pattern measures the connectivity of any POS tag followed by the verb and the second measures the connectivity of the verb and any POS tag that comes after it. The results show that the verb is slightly more connected to the left than to the right, as the IR of the two patterns is respectively .013, .020 [χ^2= 15.052, p =<.0001, d=.060, 99% CI [.975, .983]].

Given that not all words preceding a verb are the subject nor a part of it and that not all the ones following it are the object nor a part of it, we need a closer investigation. Hence, was carried out an analysis of the IR of the verb and the different syntactic structures that can play the role of subject as well as the same structures in the role of object.

As a global observation, the IR of the individual structures do not give a clear picture of the differences given their small values. For example, with personal pronouns, one of the simplest forms that a verb subject can take (1.a), the IR of the bigram (PPS VB) is equal to .002. The same IR is observed with personal pronouns used as object (2.a).

(1) a. so **I guess** all the boxcars will have …
b. **the oranges are** at Corning ...
c. **the shortest route is** via Dansville …

As for the multiword NP, like the pattern (DT NN VB) (1.b), it has an IR that is equal to .059. Concerning the subject sequence (DT JJ NN VB), as in (1.c), the IR rate is equal to .043. Similarly, the IR between the verb and object noun (VB NN) as in (2.b) is .041. Within a similar structure, but with an adjective before the noun (VB JJ NN) (2.c), the IR is equal to .023. The same goes for the sequence (VB DT NN) like in (2.d) where the IR is .050 and the sequence (VB DT JJ NN) (2.e) where IR is .028.

(2) a. no you can **carry them** both ...
b. .. we need to **get oranges** to Elmira ...
c. … we could **attach both boxcars** to one engine …
d. **wait a second** I thought well ...
e. okay **determine the maximum number** of boxcars

As for indirect complements, where a preposition is necessary to link the verb to its object, we have the trigram (VB IN NN) like in (3.a) with an IR equal to .046. While for a complement preposition phrase (VB IN DT NN) like in (3.c), the IR is .028. Besides, the IR of a verb followed by an indirect object pronoun (VB IN PPO) (3.d) is equal to 0 (we only have 12 occurrences of this pattern). Similar observations were made in the case of verbs requiring a particle (VB RP NN) (3.b), where the IR between the particle and the noun is .018 (no interruptions were observed between the verb and the particle).

(3) a. … the ones that we **filled with bananas**…

b. … **pick up oranges** for that one …
c. … as **shown on the map** ...
d. no they are already **waiting for me** …

Nevertheless, the overall interruption rate of subject structures, which is equal to .004, is about six times smaller than the interruption rate of object structures, which is .025. The difference here is statistically significant [χ^2= 54.182, p =<.0001, d=.177, 99% CI = [.974, .966]]. Please note that the effect size (*Cohen d*) cannot be big with disfluencies, given their small frequency.

3.3 Verb, particles, and prepositions

Particles are a class of invariant words that are used to change the semantics of some verbs (Malmkjaer, 2002). Their behavior is very close to the prepositions'. Some of their notable syntactic properties are worth to discuss, however. The main difference between a particle and a preposition provided in grammar manuals is that a preposition always comes before the NP. For example, it comes directly before the noun like in (4.d), before the determiner in an NP (4.e), or before a proper noun (4.f).

(4) a. ... try and **work this out** …
b. and **bring it over** to Corning …
c. … if I **drive the engine up** from Avon to Dansville …
d. so that is **from engine** E two …
e. work **at the same** time right
f. I can get **to Bath** by seven …

On the other hand, a particle can be moved around a noun, a demonstrative pronoun (4.a), object pronoun complement (4.b), or an NP complement with a determiner and a noun (4.c). In this case, the particle that behaves like a separate morpheme of the verb can be dislocated some words away from it. The hypothesis here is that all the constituents that are embedded between the verb and its particles are planned together. Please note that some verbs admit both a particle and a preposition (5.b).

During this study, eight backward patterns were identified (connections with words at the left-hand context) with 631 occurrences and eleven forward ones (connections with words to the right-hand context) with 607 occurrences. The IR of the backward patterns is 0 (out of a total of 628 cases), while the IR of the forward ones is .029. This shows that, in general, the particles play the role of an argument to a previous word rather than a predicate or argument with relation to the following word.

The data show no interruptions between the verb and the following particle (VB RP) (5.a). The difference between the general pattern XX (general bigram) with the pattern VB RP is statistically significant [χ^2= 13.389, p= <.001, d= .2328, 99% CI= [.971, .975]]. A similar pattern between the verb and the following particle and preposition (VB RP IN) is observed as in (5.b). When followed by a preposition only without a particle (VB IN), the IR is .006. Comparing this pattern to the general pattern XX gives also significant results but a smaller effect size than with VB RP [χ^2= 34.2988, p=<.001, d=.1562, 99% CI = [.971, .974]].

(5) a. to Avon to **pick up** the bananas
b. okay so it is **starting out with** a boxcar
c. I guess **by train**

To demonstrate the syntactic role of the RP after a verb, other patterns involving a verb followed by an RP have also been depicted. Interestingly, the patterns (VB RP IN) has zero interruption rate as well. As for the cases involving a verb, a particle and another POS in between, were identified two major trigrams with the categories PPO (e.g. *it*, *them*) and PRON (e.g. *this*, *those*, *that*). Besides, nine minor trigram structures involving categories such as CD (e.g. *one*), RB (e.g. *back*, *only*, *already*), and NPP (e.g. *Bath*) are also identified. These patterns have frequencies ranging from one to six cases. If we take the general pattern (VB X RP), where X is a category among the previously mentioned ones, we have a total of 135 cases with no interruptions. Compared to the general trigram pattern (XXX), this gives the following results [χ^2= 3.688, p=.054, d=.322, 99% CI[.947, .953]]. In addition, were also observed structures with fourgrams involving a determiner and a noun between the verb and its particle (VB DT NN RP). Among the nine occurrences observed in the corpus, no interruptions were observed. A recapitulation of the structures involving a verb and a particle is provided in table 2.

Structure	IR	# cases	Structure	IR	# cases
VB RP	0	534	Miscellaneous VB X RP	0	21
VB RP IN	0	45	VB DT NN RP	0	9
VB PPO RP	0	51	Total VB X RP	0	135
VB PRON RP	0	18			

Table 2 Recapitulation of the structures involving a verb and a particle

3.4 Verb's indirect objects

Some verbs in English require a preposition to introduce their object complement, called the *indirect object*. In the linguistic literature, the preposition introducing the object is processed in different ways. On the one hand, Phrase Structure Grammar (PSG) considers this preposition as a part of a complex unit, called Prepositional Phrase (PP), made of the preposition and a noun phrase. As such complex units are not allowed within the chunking framework proposed by (Abney, 1994), here, on the other hand, the preposition is given a standalone status, where it is considered to form its own chunk. Given the strong semantic correlation between the verb and its preposition, many foreign language manuals and dictionaries provide the verbs with their preposition like *depends on* and *depends to*. In this third case, the preposition has a privileged relation with the verb rather than with the noun. The IR of the bigrams (VB IN) and (IN NN) are respectively .005 and .015. This difference in frequency turns out to be statistically significant [χ^2=6.977, p=.0083, d=.101, 99% CI[.973, .995]]. Furthermore, the IR of the bigram (IN VB) is .018, which is larger than the IR of the bigram (VB IN). This difference is also statistically significant [χ^2= 9.632, p=.001, d=.103, 99% CI[.972, .990]]. This suggests that the verb is more connected with the preposition as its argument than being the argument of a preposition.

4 Discussion

4.1 Disfluencies and utterance syntactic complexity

The first question raised in this paper was whether the syntactic complexity increase yields an increase in the number of disfluencies. The reported results in table 1 confirmed this hypothesis. The correlations with the five considered measures were all positive and statistically significant. This confirms the general conception about disfluency as being caused by a heavy cognitive processing related to the task or to the linguistic complexity. For example, (Cook *et al.*, 1974) have shown that the rate of filled pauses increased with the increase of a complexity measured by the length of the following clause (no significant results were found with the subordination index devised by Frieda Goldman-Eiseler). This was also confirmed by Ferreira's work (Ferreira, 1991). A more recent work based on corpus study also shown that disfluencies occurrences correlate with the macro syntax and discourse (Beliao and Lacheret, 2013).

4.2 Verb, particle, and the planning of the complements

Given the strong relationship between the verb and its particle, this latter may be considered as a separate morpheme of the verb. Hence, an easy interpretation of the null IR between the verb and the particle in the bigram (VB RP) is to consider that this is happening because of a morphological reason; no syntax is involved here. However, similar, statistically significant, patterns were also observed with the trigram (VB X RP), where X may be any category among 11 possible complements of the verb. Although the data were not large enough to achieve significance with fourgrams, the zero IR was observed in this case as well. This is a clear indication that syntax is behind this phenomenon as it is not possible to imagine a morphological relationship between the verb and such a diverse group of categories. Put within a larger perspective, this confirms the idea that syntax is deeply embedded within the planning process of spoken language production.

4.3 Verb, its subject and object

Given the linearity of human language, it is widely thought today that language production is an incremental process. However, there are several models of incrementality that diverge in their fundamental stipulations of the timing of conceptual encoding and the timing of grammatical structures' creation. Some believe that this process is done in a "word-by-word" fashion and therefore it is completely linear (Branigan, 2008), (Kempen and Hoenkamp 1987). In other words, according to this model, during piecemeal formulation of utterances, verbs are planned only briefly before they are uttered. On the other hand, hierarchical incrementality assumes that, at the beginning of the utterance generation, its "linguistic blueprint" is formulated (Kuchinsky et al., 2011), (Zenzi and Bock, 2000). According to a lighter version of hierarchical incrementality, planning begins with the thematic structure of the event, where the relation between the agent and the patient is encoded (Bock et al, 2004). Finally, Ferreira's model of language production, which is based on Tree Adjoining Grammar, stipulates that the lexical selection of the verb is necessary before the speaker can plan the subject (Ferreira, 2000).

The data in section 3.2 show that the verb is more connected to its subject than to its object. This supports a light hierarchical incremental planning. A *verb-first* approach, such as the one proposed by (Ferreira, 2000), entails that we should not see interruptions between the verb and the subject. On the other hand, a linear incrementality would lead to equal interruption rates between the verb and its subject and its object.

4.4 Verb, particle, and preposition

The results presented in section 3.3 confirm the common conception in the classic grammar according to which particles are more tightly related to verbs than to prepositions. They also suggest, nevertheless, that prepositions have a privileged relationship with the verbs they complement. On the other hand, when the preposition is located before the verb, its IR with the verb is much larger than when it is after. This suggests that the nature of its relationship with the verb is different in this case. A possible reason is that the preposition is introducing a new proposition (via the verb) making it an important articulation point inside the sentence. One could ask if this is simply due to the prosodic structure of the utterance rather than the syntactic one.

Numerous previous studies have shown that pitch, accents, and intonation have a strong correlation with the sentence's syntactic structure (Nespor and Vogel, 2007), (Inkelas and Zec, 1990). Although several studies have attempted to use dependency grammar as a descriptive framework for prosody-syntax congruence (Mertens, 2009), (Gerdes, Hi-Yon, 2003), the majority of the existing linguistic and psycholinguistic models are based on phrase structure approaches to syntax. For example, (Cooper and Paccia-Cooper, 1980) proposed a model based on the idea that the likelihood of an intonational boundary correlates with the increase of the number of syntactic brackets at a word boundary. Hence, the likelihood of a boundary at the ends of syntactic constituents occurs more than at the beginning. Also, Ferreira (1988) proposed a model based on X-bar theory where syntax and semantics play a role in intonational phrasing. According to Ferreira this increases the semantic coherence as it minimizes the number of dependencies across units. As we saw, the patterns (VB IN) and (IN VB) have equal prosodic status (both are located at phrase borders) but different IR. This confirms that the difference is related to the nature of the syntactic relation.

5 Conclusion

This paper is about a corpus study of the interruption by simple disfluencies between key components of the utterance. The basic assumption behind this study is that interruptions depend on syntactic factors. The results confirmed some well-known facts about English syntax such as the tight interrelation between the verb and the particle. Furthermore, it also has shown a tight relation between the verb and its preposition compared to the relation between the preposition and the subsequent NP. Also, the tight relation between the verb and its subject supports the conception of light hierarchical incremental planning of language production. Beyond the direct facts, this work offers a quantitative description of the cognitive dependencies between the words with probability-like scores.

Different paths are worth to explore after this work. One of them is to study similar phenomena in language acquisition corpora. This could give us interesting insights about whether these patterns are innate or if they evolve throughout time. Covering more types of disfluencies can also bring insights about possible differences between the patterns involving each type. Finally, using a larger corpus such as the Switchboard Corpus could help confirm the obtained figures.

6 Acknowledgement

I gratefully acknowledge the helpful feedback of Professor Joseph A. Durlak on the interpretation of the statistical results.

7 Bibliography

Steven Abney. 1994. Parsing by chunks. Bell Communication Research. November 10. http://www.vinartus.net/spa/90e.pdf

Julie Beliao and Lacheret Anne. 2013. Disfluency and discursive markers: when prosody and syntax plan discourse, DiSS 2013: The 6th Workshop on Disfluency in Spontaneous Speech, Aug, Stockholm, Sweden. 54 (1), pp.5-9, 201.

K. Bock, and Levelt W.J.M. 1994. Language production. Grammatical encoding. IN M.A. Gernsbacher (Ed.). Handbook of psycholinguistics (pp.741-779). New York: Academic Press.

K. Bock, D. E Irwin and D. J. Davidson. 2004. Putting first things first. In J. M. Henderson, & F. Ferreira (Eds.), The interface of language, vision, and action: What we can learn from free-viewing eye tracking (pp. 249−278). New York: Psychology Press.

D. S. Boomer. 1965. Hesitation and grammar encoding. language and speech, Vol 8, Issue 3.

H. P. Branigan. M. J. Pickering, M. N. Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. Lingua 118, 172–189.

J. G. Carbonell and P.J HAYES. 1984. Recovery strategies for parsing extragrammatical language. American Journal of Computational linguistics. 9(3-4), pp123-146.

M. Cook, J. Smith and Lalljee, M. 1974. Filled pauses and syntactic complexity. Language and Speech. 17, 11-16.

Mark G. Core. 1999. Dialog parsing: from speech repairs to speech acts. Ph.D. dissertation, University of Rochester, New York.

William E. Cooper, Jeanne Paccia-Cooper. 1980. Syntax and Speech. Harvard University Press. ISBN 9780674283947.

G.S. Dell, Change, F., and Griffin, Z.M. (1999). Connectionist models of language production: lexical access and grammatical encoding. Cognitive Review. 23:517-542.

Fernanda Ferreira. 1988. Planning and timing in sentence production: The syntax-to-phonology conversion. Unpublished dissertation, University of Massachusetts, Amherst, MA.

Fernanda Ferreira. 2000. Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon (Ed.), Aspects of language production (pp. 291–330). London: Psychology Press.

Fernanda Ferreira. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances, Journal of Memory and Language, 30: 210-233.

Fernanda Ferreira. 2000. Syntax in language production: An approach using Tree-Adjoining Grammars. In L. Wheeldon (Ed.), Aspects of language production. Cambridge, MA: MIT Press.

Marilyn Ford and Virginia M. Holmes. 1978. Planning units and syntax in sentence production. Cognition Volume 6, Issue 1, March. Pages 35-53.

V. A. Fromkin. 1973. Speech Errors as Linguistic Evidence. The Hague, Netherlands: Mouton.

M.F. Garrett. 1980. The limits of accommodation. In V. Fromkin (Ed.), Errors in linguistic performance. (pp. 263-271). New York: Academic.

M.F. Garrett. 1988. Processes in language production. in F. J. NEWMEYER (editor), Linguistics: the Cambridge Survey, Vol. III: language: psychological and biological aspects, Cambridge: Cambridge University Press.

Kim Gerdes and Yoo Hi-Yon. 2003.The fields on the way to prosody Alternatives to phrase structure based approaches to prosody. ICPhS Barcelona, Spain, August 3-9.

A. Pearl Gordon, L. Harold Luper. 1989. Speech disfluencies in nonstutterers: Syntactic complexity and production task effects, Journal of Fluency Disorders, Volume 14, Issue 6, December Pages 429-445.

Robert Hartsuiker and Kolk Herman. 2001. Error Monitoring in Speech Production: A Computational Test of the Perceptual Loop Theory. Cognitive Psychology 42(2):113-57 April.

WO Haynes and SB Hood. 1978. Disfluency changes in children as a function of the systematic modification of linguistic complexity. Journal of Communication Disorders 1978 Feb;11(1):79-93.

Peter Heeman and James Allen. 1995. The trains 93 dialogs, TRAINS Technical note94-2, The University of Rochester computer science department, March.

Peter Heeman and James Allen. 1999. Speech repairs, intonational phrases, and discourse markers: modeling speakers' utterances in spoken dialogue, Computational Linguistics 25 (4), 527-571.

Sharon Inkelas and Draga Zec, (editors). 1990. The Phonology-Syntax Connection, The University of Chicago Press. ISBN 0226381013.

G. Kempen E. Hoenkamp. 1987. An incremental procedural grammar for sentence formulation. Cognitive Science, Volume 11, Issue 2, April, 1987 https://doi.org/10.1207/s15516709cog1102_5

Kirsten Malmkjaer. 2002. The linguistics encyclopedia, Second Edition, London/New York, Routledge.

Stefanie Kuchinsky K. Bock D. E. Irwin. 2011. Reversing the hands of time: changing the mapping from seeing to saying. Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol 37(3), May 2011, 748-756.

M. Zakaria Kurdi. 2016. Natural language processing and computational linguistics 1: speech, morphology and syntax. John Wiley & Sons, ISBN-10: 1848218486.

M. Zakaria Kurdi. 2002. Combining pattern matching and shallow parsing techniques for detecting and correcting spoken language extragrammaticalities. 2nd Workshop on RObust Methods in Analysis of Natural Language Data ROMAND 2002, Frascati-Rome, Italy - July 17.

M. Zakaria Kurdi. 2003. Contribution à l'analyse du langage oral spontané, Ph.D dissertation, Joseph Fourier University, Grenoble, France.

M. Zakaria, Kurdi. 2017. Lexical and Syntactic Features Selection for an Adaptive Reading Recommendation System Based on Text Complexity. Proceedings of the 2017 International Conference on Information System and Data Mining, Charleston, SC, USA — April 01 - 03, 2017, pages 66-69.

W. J. M. Levelt, A. Roelofs, and Meyer, A. S. 1999. A theory of lexical access in speech production. Behavioral and Brain Sciences. 22(1), 1–75.

W.J. M. Levelt. 1983. Monitoring and self-repair in speech. Cognition 14, 41-104.

Anders Lindström, Jessica Villing, Staffan Larsson, Er Seward, Cecilia Holtelius and Ab Veridict. 2008. The effect of cognitive load on disfluencies during in-vehicle spoken dialogue, In Proceedings of the 9th Annual Conference of the International Speech Communication Association (INTERSPEECH 2008), 22-26 September, Brisbane, Australia.

Kenneth Logan K., L. LaSalle. 1999. Grammatical characteristics of children's conversational utterances that contain disfluency clusters. Journal of Speech, Language, and Hearing Research, vol. 42, pp. 80–91, Feb.

Scott F. McLaughlin Walter L. Cullinan. 1989. Disfluencies, utterance length, and linguistic complexity in nonstuttering children, Journal of Fluency Disorders, Volume 14, Issue 1, February, Pages 17-36.

Marie W Meteer Ann A Taylor et al, 1995. Dysfluency annotation stylebook for the Switchboard corpus, M. Meteer and A. Taylor. Disfluency annotation stylebook for the Switchboard corpus. Department of Computer and Information Science, University of Pennsylvania. ftp://ftp.cis.upenn.edu/pub/treebank/swbd/doc/DFL-book.ps. Accessed 2018.

M. Nespor, I. Vogel. 2007. Prosodic Phonology. Berlin-New York, Mouton de Gruyter. ISBN 3110197901.

N. Nozari, G.S. Dell, M.F. Schwartz. 2011. Is comprehension the basis for error detection? A conflict-based theory of error detection in speech production. Cognitive Psychology, 63(1), 1-33.

Piet Mertens, Prosodie, syntaxe et discours : autour d’une approche prédictive, In Yoo, H-Y & Delais-Roussarie, E. (eds), proceedings d'IDP, Paris, Septembre 2009, ISSN 2114-7612, pp. 19-32, 2009.

Ralph Rose. 2017. Silent and filled pauses and speech planning in first and second language production. In Eklund, R. and Rose, R. (Eds.), Proceedings of DiSS 2017, Disfluency in Spontaneous Speech. Stockholm, Sweden: Royal Institute of Technology (KTH), ISSN 1104-5787, pp. 49-52.

Michael Wagner. 2016. Information Structure and Production Planning. In Caroline Fery and Shin Ishihara editors, Oxford Handbook on Information Structure. Oxford University Press, Oxford.

M. Griffin Zenzi and Kathryn Bock. 2000. What the Eyes Say About Speaking. Psychological Science Vol 11, Issue 4, 2000, 274-279.