Frozen by context: Pragmatic factors of syntactic freezing

GOUMING MARTENS (iD)

Abstract

Syntactic freezing has mainly been approached from a structural point of view, recently though, more cognitive approaches in terms of processing costs have been proposed. One such processing account is the additive account. According to this approach, the freezing effect is best explained as an additive effect of two syntactic processes coming together, rather than being a phenomenon on its own. Another processing account argues that the freezing effect is the result of a prosodic garden path according to which extraction can only take place from a prosodically focused constituent. The current study provides empirical evidence for a less discussed factor contributing to the freezing effect, namely a pragmatic one. The pragmatic account requires frozen sentences to have contextually given referents. If no such referent is present, the sentence becomes less acceptable. The need for such a referent comes from the non-default word order associated with frozen sentences, which often highlights/focuses a certain constituent. Several experiments were run to test the pragmatic account. Based on the results it was concluded that pragmatic factors play a significant role in explaining the apparent freezing effects. Other factors however, seem to contribute to this effect as well since this effect cannot be fully explained in terms of pragmatic factors solely.

CORRESPONDING AUTHOR:
Gouming Martens
McGill University, CA
gouming.martens@mail. mcgill.ca

1 Introduction

Freezing is a syntactic phenomenon according to which a constituent that has been moved from its base-generated position becomes a barrier for extraction, meaning that no material inside the 'frozen' constituent can be extracted out of it (Corver 2006).

An illustration of this phenomenon is provided in Example (1). Whenever a DP is in its base-generated position (1a), extraction of a wh-phrase out of this DP is considered grammatical (1b): The wh-phrase *Whose books$_i$* is extracted out of a base-generated DP *reviews of t$_i$* which is considered grammatical. On the other hand, when that same DP (i.e. *reviews of t$_i$* has been moved out of its original position (1c), extraction of a wh-phrase out of this DP is no longer considered grammatical (1d).[1] The current study takes a closer look at frozen constructions containing fronted DPs like the ones in Example (1).

(1)　　　(Corver 2006: 1)
　　　a.　　I think that John never reads [reviews of his books]
　　　b.　　Whose books$_i$ do you think that John never reads [reviews of t$_i$]?
　　　c.　　I think that [reviews of his books]$_i$ John never reads t$_i$
　　　d.　　*Whose books$_i$ do you think that [reviews of t$_i$]$_j$ John never reads t$_j$

The idea that there are certain syntactic domains that prevent extraction of syntactic material (i.e. syntactic islands), has been discussed ever since the early days of generative grammar (Chomsky 1964; Ross 1967). It was Ross (1967) however, who first mentioned the term *freezing* in his dissertation. In his dissertation it was argued that there is a syntactic constraint that prevents extraction out of a frozen constituent (Frozen Structure Constraint FSC). Ever since Ross's seminal work, many different syntactic analyses have been proposed (Wexler & Culicover 1980; Müller 2010).

The idea that syntax is solely responsible for the freezing effect, has long been the standard assumption. More recently though, other factors involving processing difficulties have been argued to play an important role in explaining the freezing effect as well (Hofmeister et al. 2015; Winkler et al. 2016; Konietzko et al. 2018). Several studies in favor of a processing account have argued that the freezing effect is the result of two syntactic processes occurring simultaneously, namely movement and extraction (Hofmeister et al. 2015; Konietzko et al. 2018): According to these so-called *additivity* accounts, there is no such phenomenon as syntactic freezing. The current study investigates this claim for frozen sentences involving a fronted DP.

So far, there is no consensus on how freezing should be characterized and which factors play a role. As already pointed out by Culicover & Winkler (2010), different freezing constructions require different explanations and therefore there is not just one account that would be correct. The main goal of this study is to provide empirical evidence for a new factor that plays a role in explaining the freezing effect, namely a pragmatic factor. In this study it will be argued that context plays an important role in (partially) explaining the freezing effects as observed in sentences with fronted DPs. In addition, it will become clear that in the case of frozen sentences with a fronted DP, a pragmatic account of freezing is preferred over a prosodic account. Important to note is that the aim of this study is not to propose a completely new analysis nor argue for one specific analysis, rather the aim is to provide empirical evidence showing that pragmatic factors play a significant role in explaining the freezing effect.

1.1 Structural accounts

Ross (1967) observed that no element of an extraposed relative clause can be relativized or questioned, as becomes clear from Example (2). In Example (2a) the relative clause *who had worn this coat* has been extraposed from a position adjacent to the lexical head *A girl* to a position following the verb *came in*. The relative pronoun *which* cannot be moved out of the extraposed relative clause and move to the left periphery of the larger relative clause that modifies the antecedent *the coat* (2b).

1　　Although these "grammaticality" judgments have been generally accepted, it will become clear in the current study that such judgments are less categorical than often presented. From now on, the term *acceptability* will be used rather than *grammaticality* to mark the non-categorical nature of native speakers' intuitions. Though not discussed in this paper, there has been a lot of debate about grammaticality and acceptability judgments. For those interested, I would like to refer to Keller (2000); Schütze (2016); Sprouse (2007).

(2)   (Ross 1967: 294–295)
    a.   A girl t_j came in [who had worn this coat]$_j$.
    b.   *[The coat [which$_j$ [a girl t_i] came in [who had worn t_j]$_i$]] was torn.

To account for the ill-formedness of Example (2b), Ross (1967) formulated the Frozen Structure Constraint (FSC).

(3)   *Frozen Structure Constraint* (Ross 1967: 295):
    "If a clause has been extraposed from a noun phrase whose head noun is lexical, this noun phrase may not be moved, nor may any element of the clause be moved out of that clause."

In a later paper, Ross incorporated the FSC under a more general island constraint: The Immediate Self Domination Principle (ISF). According to this principle, no element may be chopped out of a node which immediately dominates another node of the same type (Ross 1974). Given that self-dominating nodes represent Chomsky-adjunction configurations, the ISP states that extraction of a constituent out of an adjoined structure is prohibited: This means that in a structure like [$_A$ A B], nothing can be extracted out of B. As implied by the ISP, islandhood not only holds for base-generated adjunction configurations, but also for movement derived adjunction configurations. Going back to Example (2) (partially repeated in (4) it becomes clear how the ill-formedness of (2b) can be explained in terms of the ISF. In (4), the relative clause *who had worn this coat* is immediately dominated by a node S which is of the same type as the constituent *A girl came in.* Therefore no element can be extracted out of the relative clause according to the ISF. This would explain the ill-formedness of Example (2b).

(4)   [$_S$ [$_S$ A girl t_i came in] [$_{RC}$ who had worn this coat]$_i$].

Building on Ross (1967) and Ross (1974), Wexler & Culicover (1980) proposed their own version of a Freezing Principle to account for the data which was based on considerations of language learnability. The basic idea of their principle is that any structure created via a syntactic transformation that is not compatible with the original base phrase structure rules of that language, is frozen. Such a derivation is called "non-structure-preserving", in the sense of Emonds (1970; 1976).

Example (5a) illustrates a ditransitive sentence. For this sentence the base phrase structure rule is VP > V NP PP. Since no syntactic transformation has applied to the PP it is not frozen and therefore extraction of a wh-word is possible as is illustrated by the acceptability of (5b). In Example (5c) on the other hand, heavy NP shift has taken place, in other words a syntactic transformation has applied to sentence (5a) which moves the NP to a position following the PP. The sentence created after transformation is no longer structure-preserving (i.e. cannot be derived from the base generated phrase structure rules which would be VP > V DP PP) and therefore any element within this non-structure-preserving node is frozen. For this very reason one cannot extract a wh-phrase out of the PP as is illustrated in (5d).

(5)   (Wexler & Culicover 1980: 279)
    a.   John gave [the picture that was hanging on the wall] [to Bill].
    b.   Who$_i$ did John give [the picture that was hanging on the wall] [to t_i]?
    c.   John gave [to Bill] [the picture that was hanging on the wall].
    d.   *Who$_i$ did John give [to t_i] [the picture that was hanging on the wall]?

Another structural account by Müller (2010) explains freezing in terms of phases and the Phase Impenetrability Condition (PIC). The main idea behind this analysis is that a phrase which occupies the specifier of a phase head is frozen (i.e. cannot be extracted from), only if it is the last-merged specifier within that phase. This is defined in Müller's version of CED, see (6) below.

(6)   *Condition on Extraction Domain* (Müller 2010: 38)
    a.   Movement must not cross a barrier
    b. α is a barrier if the operation that has merged α in a phase Γ is the final operation in Γ.

It should become clearer from Example (7) how the CED would account for frozen sentences. In (7a), extraction of the wh-word is possible as the verb *was* moves via the edge position of DP to the edge of VP and finally ends up in specCP. This mechanism is based on the idea that only elements that are located at the edge of a phase can move out of a phase as is defined by the PIC (Chomsky 2001). Crucially for Müller (2010) all phrases are phases: This is in line with Koster (1978) and Van Riemsdijk (1978). In (7b), on the other hand the object has been scrambled to the (outer) specifier of VP, which makes it the last-merged element in the VP phase. As a consequence the phase head V has become inert and the DP is rendered frozen. This means that no element can move out of the frozen phrase.

(7)   *German*

　　a.　　Was$_j$ hat Otto [$_{VP}$ immer [t$_j$ für Romane] gelesen]?
　　　　　what has Otto 　　always t$_j$ for novels 　read
　　　　　'What kind of novels has Otto always read?'

　　b.　　*Was$_j$ hat Otto [$_{VP}$ [t$_j$ für Romane]$_i$ immer t$_i$ gelesen]?
　　　　　what has Otto 　　t$_j$ for novels 　　always t$_i$ read
　　　　　'What kind of novels has Otto always read?'

So far, Müller's CED seems to be yet another way of describing the freezing phenomenon and it would therefore not differ from any other structural analyses like Ross's freezing principle or Wexler-Culicover's freezing principle. What sets Müller's CED account apart from the other accounts is the prediction it makes with regard to "repairing" a frozen sentence. According to this account a frozen sentence can be repaired by scrambling something over it, which makes the frozen constituent no longer a barrier for further movement. This is called *melting*, as is illustrated in example (8) below.

According to Müller, *was für Bücher* in (8a) is frozen, because it is the last-merged in the specifier-position of vP and therefore forms a barrier for movement. However, based on Müller's analysis it is predicted that the sentence should become grammatical, whenever this frozen constituent is no longer the last-merged specifier. This is done in Example (8b) by scrambling *den Fritz* over *was für Bücher*. In this case *was für Bücher* is no longer the last-merged element and therefore the sentence is acceptable.

(8)   *German* (Müller 2010: 61)

　　a.　　*Was$_1$ haben [$_{DP3}$ t$_1$ für Bücher] 　　[$_{DP2}$ den 　　Fritz] beeindruckt?
　　　　　what have 　　t$_1$ for books.NOM 　　the.ACC Fritz impressed
　　　　　'What kind of books impressed Fritz?'

　　b.　　Was$_1$ haben [$_{DP2}$ den 　　Fritz] [$_{DP3}$ t$_1$ für Bücher] 　　beeindruckt?
　　　　　what have 　　　the.ACC Fritz 　　t$_1$ for books.NOM impressed
　　　　　'What kind of books impressed Fritz?'

As will become clear in the next section, rather than adopting a syntactic account for melting like Müller (2010) did, Culicover & Winkler (2010) explain this phenomenon in terms of focus patterns and prosodic constraints.

So far, we have seen different accounts that try to capture the freezing phenomenon in terms of syntactic principles. The only thing that all structural accounts seem to agree on is that sentences involving wh-extraction from a moved constituent, i.e. so called frozen sentences, are less acceptable than sentences involving wh-extraction from a based-generated constituent. The current study makes an attempt at figuring out what causes this lower acceptability of frozen sentences.

More recently, attempts have been made to account for freezing effects in terms of a constraint on language processing rather than a constraint on syntax as we will see in the following sections.

1.2 Additivity effects

Several studies in favor of a processing account have argued that that freezing effects are the result of two syntactic processes occurring simultaneously, namely movement and extraction (Hofmeister et al. 2015; Konietzko et al. 2018). Each syntactic process lowers the acceptability

judgments and adding up these lowered acceptability judgments would result in a net unacceptability rating giving rise to an apparent freezing effect. In other words, freezing is an additive effect of two costly syntactic processes occurring simultaneously.

Hofmeister et al. (2015) tested this idea experimentally by constructing sentences containing extraposed PPs, see Example (9). The stimuli were presented to participants in an online acceptability judgment task using Mechanical Turk.

(9) a. Kenneth revealed which president he overheard a nasty remark about earlier.
         (+ EXTRACTION, – EXTRAPOSITION)
     b. Kenneth revealed which President he overheard a nasty remark earlier about.
         (+ EXTRACTION, + EXTRAPOSITION)
     c. Kenneth revealed that he overheard a nasty remark about the President earlier.
         (– EXTRACTION, – EXTRAPOSITION)
     d. Kenneth revealed that he overheard a nasty remark earlier about the President.
         (– EXTRACTION, + EXTRAPOSITION)

The results in *Figure 1* show that both wh-extraction and PP extraposition lower the acceptability judgement. There also seems to be a slight interaction effect which suggests that PP extraposition lowers the acceptability score more when there is no wh-extraction than when there is wh-extraction. The statistical analysis confirms that there is a main effect of extraction and extraposition, but crucially there is no significant interaction effect.
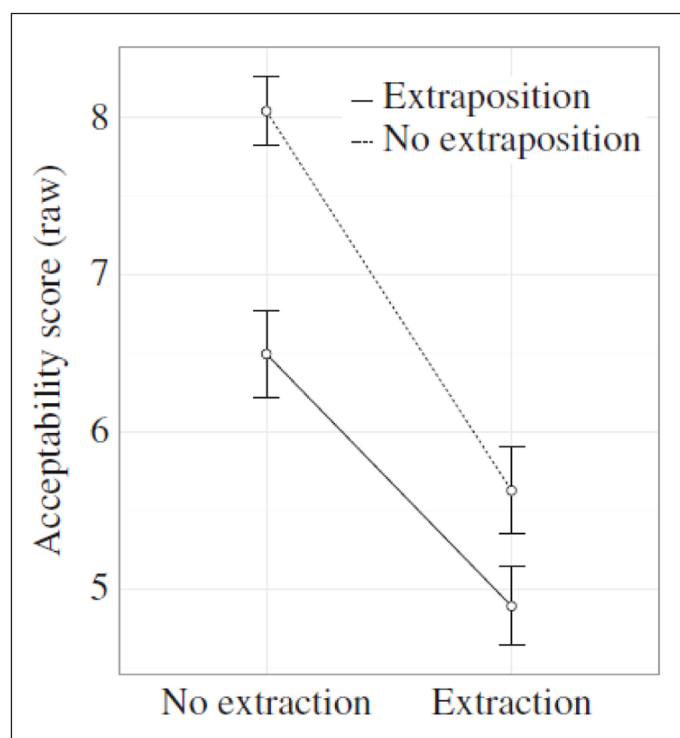
Figure 1 Acceptability scores of frozen sentences containing extraposed PPs.

Based on these results the conclusion was drawn that the effect of each variable on the acceptability judgments is sufficient to account for the net unacceptability of freezing sentences. If freezing was truly a phenomenon on its own, one would have expected an interaction effect between the two variables: This would indicate that one cannot simply add up the effects of extraction and extraposition to explain the freezing effect. Since no significant interaction effect was found, it was concluded that freezing is an additive effect of movement and extraction.

Based on the fact that previous studies have linked both these displacement phenomena (i.e. extraction and extraposition) to increased processing costs, Hofmeister et al. (2015) indirectly conclude that the apparent freezing effect is best described in terms of processing costs. It is important to notice that no direct evidence for this claim was provided by Hofmeister et al. (2015).

Similar results were found by Konietzko et al. (2018) who looked into frozen sentences involving heavy NP shift. From their results they also concluded that the freezing effect can be best

explained in terms of additivity effects: the unacceptability of frozen sentences can be predicted from the independent penalties attributed to heavy NP shift and extraction out of the heavy NP.

In line with the previous studies, the current research sets out to test the additivity hypothesis for yet another case of freezing, namely sentences involving DP-fronting (see Example (10). Experiment 1.1 tests the additivity hypothesis to see whether the freezing effect is best described as a phenomenon on its own or as the additive result of two separate processes coming together.

(10)   [Which students]$_i$ do you think [the names of t$_i$]$_j$, she always forgets t$_j$.

The additivity account would predict a main effect of wh-extraction and *DP*-fronting but no interaction between the two. If there is a significant interaction effect, then this would indicate that the results cannot simply be predicted based on the sole contribution of each syntactic process. Therefore a significant interaction effect would support the idea that freezing is a phenomenon on its own. Based on Hofmeister et al. (2015) and Konietzko et al. (2018) however, it can be expected that a similar additive effect will be found for frozen sentences involving DP-fronting when presented out of the blue.

It should be noted that the additivity account itself does not directly provide evidence for a processing account or a structural account. Rather it tells us whether the freezing effect is a phenomenon on its own or whether it is the result of two syntactic processes happening simultaneously. The additivity account itself does not tell us what factors (be it structural, processing or yet another factor) would best explain the results.

1.3 Garden paths

A different approach is taken by Culicover & Winkler (2010), according to them there are two different cases of processing failures that can be distinguished in relation to the freezing effect. On the one hand there are cases where the ill-formedness is due to a processing failure in which the processor posits a gap, following a so-called *active filler strategy* to form a chain with the A' constituent (Crain & Fodor 1985; Frazier 1987; Hofmeister & Sag 2010). When the postposed NP with the actual gap is processed, the initial hypothesis has to be revised and the chain has to be reconstructed, which produces the freezing effect. The freezing cases that Hofmeister et al. (2015) try to account for are of this nature and are termed *chain garden paths* by Culicover & Winkler (2010), as illustrated in (11).

(11)   Kenneth revealed [which President]$_i$ he overheard $_{–i}$ a nasty remark earlier about t$_i$.

The garden path is created by a potential gap immediately following the verb *overheard* in which *which president* could be interpreted. However as soon as one encounters the DP *a nasty remark*, the parser has to correct the earlier prediction and needs to place the gap somewhere else. This misinterpretation of the parser is what creates a chain garden path and would explain lower acceptability ratings.

1.3.1 Freezing as a prosodic garden path

On the other hand, there are freezing cases where the ill-formedness is due to a prosodic garden path rather than a chain garden path. The main idea here is that the lower acceptability of freezing sentences is due to a mismatch between the default intonation pattern and its syntactic form. In this case the freezing examples violate a prosodic constraint which states that extraction can only take place from a focused constituent (Bayer 2005; Winkler et al. 2016). Culicover & Winkler (2010) posit two hypotheses that are relevant to their prosodic-based analysis:

(12)   *Prosodic Garden Path Hypothesis (PGH)* (Culicover & Winkler 2010: 15):
         "When parsing decontextualized linguistic examples, the processor projects the
         default prosodic pattern of the particular language and construction onto the target
         sentence. In freezing configurations, the default intonational pattern leads to a
         prosodic garden path defined as a prosody-information structure conflict. Reanalysis
         triggers the search for a context-dependent non-default prosodic pattern."

(13)   *Information Structure Hypothesis (ISH)* (Culicover & Winkler 2010: 15):
     "Freezing effects which result from extraction out of non-canonical constructions can be repaired by discourse appropriate non-default prosody."

Let us look at an example to get a better idea of the PGH and ISH. We will take example (14) from the previous section to show how the same example can be explained in terms of a prosodic constraint rather than a syntactic constraint. Remember from the previous section that according to Müller (2010) Example (14a) is ungrammatical because *für Bücher* is located on the edge of VP. For Culicover & Winkler (2010), the lower acceptability arises because sentence (14a) deviates from the default intonation pattern. The default intonation pattern (for German in this case) would be one where the pitch accent is assigned to the preverbal DP, according to Culicover & Winkler (2010). In case of (14b) *Bücher* gets accented which is expected since it is new information (15a).

(14)   *German* (Müller 2010: 61)

     a.   *Was$_1$ haben [$_{DP3}$ t$_1$ für Bücher]   [$_{DP2}$ den   Fritz] beeindruckt?
         what have   t$_1$ for books.NOM   the.ACC Fritz   impressed
         'what kind of books impressed Fritz'

     b.   Was$_1$ haben [$_{DP2}$ den   Fritz] [$_{DP3}$ t$_1$ für Bücher]   t$_2$ beeindruckt?
         what have   the.ACC Fritz   t$_1$ for books.NOM t$_2$ impressed
         'What kind of books impressed Fritz?'

(15)   a.   Was haben [$_{DP2}$ den Fritz] [$_{DP3}$ t$_1$ für BÜCHER] t$_2$ beeindruckt ?
                       L   H*L   H%
         'What kind of BOOKS have impressed Fritz?'

     b.   *Was haben [$_{DP3}$ t$_1$ für Bücher] [$_{DP2}$ den FRITZ] beeindruckt ?
                       L   H*L   H%
         'What kind of books have impressed FRITZ?'

Applying a default intonation pattern to sentence (14a) results in the sentence accent being assigned to the discourse given DP *den Fritz*(15b). This is what causes the prosodic garden path effect according to Culicover & Winkler (2010).

There are two conflicting information structural mismatches that occur during the parsing of (15b). First of all, it is argued that the *was-für* extraction in (15b) applies to the topic constituent *für Bücher*, which violates the requirement that *was-für* extraction is preferred from a focus constituent. In general, topic constituents are said to be prosodically unmarked.[2] Assuming a default prosody then, the nominative constituent *für Bücher* remains prosodically unmarked.

Second, the processor projects a nuclear pitch accent onto the DP *den Fritz* in preverbal position, which violates the default implicit prosody according to which definite DPs tend to be discourse given and typically remain unaccented. The observed prosody-information structure conflict can be characterized in the following way: the default focus position of the sentence is occupied by the discourse given unfocused DP *den Fritz* and the target constituent of extraction *für Bücher*, which needs to be focused, is in an aboutness topic position.

To resolve this prosodic garden path one has to change the intonation pattern such that the correct elements receive the correct accents. For Example (15b) that would mean assigning an non-default intonation pattern in which *für Bücher* receives sentence pitch accent and the DP *den Fritz* is unaccented, see Example (16a).[3]

(16)   a.   Was haben [$_{DP3}$ t$_1$ für BÜCHER.NOM] [$_{DP2}$ den.ACC Fritz] beeindruckt ?
                       H*L   L   H%
         'Which BOOKS have impressed Fritz?'

---

2   Note here that an assumption needs to be made in terms of the order of syntactic processes. The assumption here is that wh-extraction happens after the DP has moved to topic position, otherwise there would be no violation.

3   There is some disagreement among native speakers whether this is actually acceptable.

By means of a prosodic constraint, Culicover & Winkler (2010) have provided an alternative explanation as to why freezing sentences are perceived to be ill-formed when no context is given. The fact that these sentences can be saved by applying a non-default intonation pattern indicates that one is not dealing with a purely structural constraint. The basic idea is that extraction out of a constituent is only possible if that constituent has been prosodically focused (Bayer 2005; Winkler et al. 2016).

1.3.2 Winkler et al. (2016)

Winkler et al. (2016) put the prosodic account to the test. They took a *constraint satisfaction* approach to explain freezing. According to constraint satisfaction theories, sentence processing involves several possible sentence structures that are simultaneously activated but that have different levels of activation. The sentence structure with the highest activation level, i.e. the one that is most likely to be correct, wins the competition for activation over the other possible analyses. In case of frozen sentences, Winkler et al. (2016) argue that the most highly activated prosodic structure conflicts with requirements of an information-structural constraint whose application is required by the syntactic form.

In line with Culicover & Winkler (2010), they propose that by manipulating the context (in this case a contrastive focus context) a non-default intonation can be induced. Consequently, the non-default contrastive focus contour can repair the construction by supporting the structure analysis that matches the syntactic structure.

Their stimuli consisted of target sentences either preceded by a context invoking a default focus (17) or a context with contrastive focus on the subject (18). The discourse particle *denn* marks the boundary between the topic and the focus field (Bayer & Obenauer 2011; Bayer 2012). The default focus context indicates that the target sentence contains new information. In that case the default accent would be assigned to *den Minister*, leaving the subject *Ärzte* unaccented and therefore wh-extraction would be less acceptable. In the contrastive focus context however, the context is supposedly forcing a contrastive accent on the subject, thus making wh-extraction acceptable.

(17)   (default focus)
         Sag mal:
         'Tell me:'
         a.   Was haben den für Ärzte   den den   MINISTER kritisiert?
             what have PRT for doctors.NOM PRT the.ACC minister criticized?
             'What kind of doctors criticized the MINISTER?'

(18)   (contrastive focus on subject)
         Dass den Minister Journalisten kritisiert haben, weiß ich schon, aber
         'I already know that journalists criticized the minister, but'
         a.   Was haben den für ÄRZTE   den den   Minister kritisiert?
             what have PRT for doctors.NOM PRT the.ACC minister criticized?
             'What kind of DOCTORS criticized the minister?'

Their results confirm that a contrastive context significantly improves a frozen sentence, at least for German *was für* constructions. Based on this it was concluded that freezing is best described in terms of a prosodic constraint. However it should be noted that it does not directly follow from the results that the prosodic analysis is the only viable analysis. In fact, the results by Winkler et al. (2016) contain a confound and could be interpreted differently: The confound here is that the comparison between default focus in (17) and contrastive focus on the subject (18), is not actually testing the claim whether extraction can only take place from a focused element or not. Rather it is testing the effect of having a contrastive context versus having an all new information context. To specifically test for the *extraction-under-focus* claim, one has to add another contrastive focus context in which the contrastive focus is not on the moved constituent but on another constituent in the sentence. The prosodic account would predict that a sentence is only acceptable when there is contrastive focus on the moved constituent and not when the contrastive focus is on any other element in the sentence. The current research adds such a contrastive focus context and it will become clear that the results do not support a prosodic account. Rather they seem to point towards a pragmatic explanation of the data.

1.4 Freezing as a pragmatic phenomenon

In this section, we will explore how the freezing effect can be explained in terms of pragmatic requirements.

Birner & Ward (1998) argue that fronted constituents in English can be seen as representing "information standing in a contextually licensed partially ordered set relationship with information invoked in or inferrable from the prior context". Although there are different views on what the exact function of DP-fronting is and how many different kinds of fronting constructions there are, most accounts seem to agree that there is at least one requirement or constraint which seems to apply to all DP-fronting constructions: All DP-fronting constructions require the referent or denotation of the fronted constituent to be anaphorically linked to the preceding discourse (Reinhart 1981; Vallduví 1993; Ward & Birner 1994).

This requirement will play an important role in the experiments reported in this article. As already suggested by Culicover & Winkler (2010) and Winkler et al. (2016), frozen sentences will improve by providing a contrastive context. According to them, the improvement of frozen sentences should be attributed to a change in prosody induced by a contrastive context. Unlike the prosodic accounts, the current study argues that the reason that these frozen sentences improve is because of the presence of a contextual anchor/referent with which the fronted DP can associate. Let's look closer into why a contextual anchor would improve frozen sentences.

1.4.1 Topic/Focus

DP-fronting is often associated with topicalization. In English two types of topicalization are distinguished in general: *focus topicalization* and *topic topicalization* (Gundel 1977; Prince 1981; Ward et al. 2002; Ward & Birner 2004). These are referred to as "Focus Movement" and "Topicalization" by Prince, and as "Focus Preposing" and "Topicalization" by Ward respectively. In focus topicalization, the preposed constituent is assigned the main accent (indicated by the underlining in Example (19). In this example, John is new information and is interpreted as focal in this sentence.

(19)   (Gundel 1977: 134)
         a.   Who did he call?
         b.   JOHN he called.

In topic topicalization, the main pitch accent is placed on the focused element which carries new information and does not correspond to the topicalized element of the sentence (Prince 1998). We will see that some stimuli would correspond to a topic topicalization and others to a focus topicalization in the experiments presented in this paper.

(20)   (Gundel 1977: 134)
         a.   What about John?
         b.   John, he CALLED.

In topicalization constructions, the preposed element provides an anaphoric link to the preceding discourse (Gregory & Michaelis 2001; Ward & Birner 2004). Therefore, topicalization contributes to the cohesion of a text (Hietaranta 1984).

A lot of confusion arises when dealing with topic and focus as already became clear from the different terminology used by different authors for the same phenomenon. The topic-focus distinction has been widely associated with the division between given and new information in a sentence. There has been disagreement and confusion, however, with regard to the exact nature of this association. According to Gundel (1988); Gundel & Fretheim (2004), part of the confusion is a result of conflating two types of givenness/newness: these are referential givenness-newness and relational givenness-newness.

Referential givenness-newness involves a relation between a linguistic expression and a corresponding non-linguistic entity in the speaker/hearer's mind, the discourse (model), or some real or possible world, depending on where the referents or corresponding meanings of these linguistic expressions are assumed to reside. We will see that all the stimuli in the first sub-experiment (1.1) are referentially new whereas in sub-experiment (1.2) they are referentially given in the sense that the fronted DPs have been previously mentioned in the context.

Relational givenness-newness, by contrast, involves a partition of the semantic/conceptual representation of a sentence into two complementary parts, X and Y, where X is what the sentence is about (the logical/psychological subject) and Y is what is predicated about X (the logical/psychological predicate). Relational givenness can be associated with topic/focus. Topics are relationally given, by definition, in the sense that they are what the sentence/utterance is about.

That does not mean that a certain constituent cannot be both a topic and receive focus as claimed by Rizzi (1997). Gundel & Fretheim (2004) argue that a similar confusion has arisen between two conceptually distinct notions of focus: "on the one hand there is a relational notion of focus – the information predicated about the topic; on the other hand there is a referential focus – material which the speaker calls to the addressee's attention, thereby often evoking a contrast with other entities that might fill the same position." These have been called information focus and contrastive focus, respectively. Although by definition, topic and information focus are mutually exclusive, it is possible for a constituent to be both a topic and receive contrastive focus if the speaker wants to direct an addressee's attention to the topic.

In the current research we will see that referential givenness plays a more important role in assessing freezing cases than prosodic/contrastive focus does (as would be predicted by the prosodic constraint analysis).

1.4.2 Alternative propositions

According to Rooth (1985) one of the primary functions of focus is evoking alternatives. Taking it even one step further, Büring (1999) claims that both topic, or Sentential topic (S-topic), and focus invoke alternative propositions. Whichever element in the sentence receives focus will be part of the set of propositions that are considered well-formed alternatives. This is what Büring (1999) calls the *Focus value*. Following Rooth (1985)'s idea, this can be formalized by deriving from a sentence S a second semantic value, its Focus value, $[[S]]^f$ for short. If the Focus is on the object NP, $[[S]]^f$ is the set of propositions one gets by sticking in alternatives for the Focus.

The S-topic is very similar to Focus in a way that they both induce alternatives. These alternatives though, do not have any impact on the Focus value, instead, it is a set containing different Focus values (Topic Value), i.e. a set of sets of propositions. Regardless of whether DP-fronting functions as topic or focus, or even both, the fact that both topic and focus invoke alternatives could potentially explain why sentences with preposed DPs would become more acceptable when such an alternative is already present in the preceding context.

The goal of the current study is not to determine whether the fronted DP is topic or focus nor is it to choose one freezing account over another one, rather it proposes a previously not considered factor that contributes to the freezing effect. Pragmatic factors do not rule out a processing or structural explanation of freezing. In fact, the freezing effect is most likely the result of various factors. Crucially though, it will be shown that pragmatic factors play an important role in explaining the freezing effect.

1.4.3 D-linking

All stimuli in this study consist of so called D(iscourse)-linked wh-words. These are wh-words like *which...* as opposed to non D-linked wh-words like *what* or *who*. One known observation about D-linked wh-words is their insensitivity to syntactic islands, compare for instance the Examples in (21) below.

(21)   (Boeckx & Grohmann 2004: 5)
    a.   *Who$_i$ did John wonder whether Mary loved t$_i$?
    b.   ?[Which man]$_i$ did John wonder whether Mary loved t$_i$?

The operation moving the D-linked wh-word *which man* out of the embedded clause and across an island boundary does not need to be subject to the same locality conditions as regular A'-movement. It is beyond the scope of this paper to go into more detail on how to formalize this distinction.

In addition, as the name already suggests, these D-linked elements are somehow associated with the discourse. A wh-question like in (22b) needs some discourse referent and cannot be

used out of the blue, like the question in (22a). This discourse linkage seems to be somewhat similar to the way a topic is said to associate with a discourse. Similarly then the expectation is that when an appropriate discourse is provided, sentences containing a D-linked wh-word would improve compared to when no such discourse is provided.

(22)   (Boeckx & Grohmann 2004: 4)
        John bought something expensive yesterday.
        a.   What did he buy?
        b.   #Which car did he buy?

2 Expectations

In this section I will recapitulate what has been discussed in the previous literature, highlight what is of interest for the current study and make predictions about the expected results.

As becomes clear from the previous literature, sentences involving wh-extraction from a moved constituent, i.e. frozen sentences, are less acceptable than sentences involving either wh-extraction or movement. In addition, there is evidence that the individual lowering effects of extraction and movement can be added up to account for the freezing effect. In terms of statistics that would mean that a significant main effect of wh-extraction and DP-fronting is to be expected, but no interaction effect.[4] This claim has been supported by Hofmeister et al. (2015) who tested for extraposed PPs and by Konietzko et al. (2018) who examined the heavy NP shift. The current research looks into sentences involving DP-fronting, a similar additive effect is expected to be found: That means that a significant main effect of wh-extraction and DP-fronting is expected, but no interaction effect. As already made clear earlier, finding an additive effect still says little about which factors contribute to the freezing effect and in fact additivity can be explained by different accounts.

In the previous section it has also been observed that some cases of freezing are not completely unacceptable and can be improved: Think of Müller (2010)'s melting examples and Culicover & Winkler (2010)'s prosodic garden paths. The focus in this study will be on the prosodic account. The current study compares the prosodic garden path analysis against a pragmatic analysis. According to the prosodic account, extraction out of a constituent is only possible if that constituent has been prosodically focused. The expectation is that when a moved constituent is not prosodically focused, the sentence should become less acceptable.

A pragmatic account on the other hand, only requires a referential anchor in the previous context to which the moved constituent can refer. That means that regardless of whether or not that moved constituent is prosodically focused, as long as there is a contextual anchor present, the sentence should be more acceptable than without such referent.

The current study argues for a pragmatic account instead of a prosodic account of freezing (at least in the case of DP-fronting). The expectation is that there will be no significant effect of prosodic focus, but there will be a significant effect of the presence of a contextual anchor. In order to compare the prosodic account against the pragmatic account, production data will be gathered as well. If the prosodic account is on the right track the expectation is that wh-extraction is acceptable only when the moved constituent is focused. For the pragmatic account the prediction is that there is no correlation between the acceptability of a sentence and the prosodic focus in the sentence. It is important to note that the current study will not propose a complete new pragmatic analysis nor argue for one specific analysis, the goal is merely to provide empirical evidence that pragmatic factors play a major role in explaining this specific freezing phenomenon.

3 Experiment 1

Experiment 1 tests the hypothesis that freezing cases are accommodated by having a contextually given referents to the fronted and contrastive DP. Moreover, Experiment 1 will provide evidence

---

4   An interaction effect would indicate that there are extraneous effects that cannot be solely explained in terms of the individual variables.

against a prosodic account of freezing as formulated by Culicover & Winkler (2010). In order to test for a prosodic account, a condition has been created in which the contrastive focus is not on the fronted DP but on the main verb. If the prosodic account is correct it is expected that only in case of contrastive *DPFocus*, frozen sentences would become more acceptable, but not when the main verb is focused. According to the pragmatic account on the other hand, there should be no difference between the two contrastive conditions.

In addition, participants are asked to produce the target stimuli in order to clarify which intonation pattern they have in mind when assessing the sentences. This experiment consists of two sub-experiments. The first sub-experiment (Experiment 1.1) consists of stimuli sentences without any context and Experiment 1.2 consists of stimuli sentences with a context introducing a referent.

The first sub-experiment sets out to test the additivity account to determine whether freezing is a phenomenon on its own or whether it can be derived from the joint effects of wh-extraction and DP-fronting. The second sub-experiment will tell us whether the freezing effect can be completely accounted for based on the presence of a contextually given referents. The results of these two sub-experiments will be combined to see the contribution of context/given-referent to the sentences containing fronting and wh-extraction.

3.1 Experiment 1.1

Experiment 1.1 tests the effects of DP-fronting and wh-extraction on the acceptability of sentences presented out-of-the-blue, i.e. sentences without any context. This experiment by itself does not tell us which factors (i.e., syntactic, prosodic and/or pragmatic) have an effect on freezing, all it tests is the additivity account.

If freezing is a phenomenon on its own, one expects to find a significant interaction effect between fronting and extraction according to which the difference in fronting is larger in the *Extraction* condition than in the *No Extraction*-condition. In other words, we expect to find a super-additive effect, i.e. the effect of the combination of the two syntactic processes is larger than would have been predicted based on their individual contributions. If on the other hand, the additive account is correct then one would expect to find a main effect of *extraction* and *fronting*, but no interaction effect between the two variables. Based on Hofmeister et al. (2015) and Konietzko et al. (2018) we expect to find a significant main effect of fronting and extraction, but no interaction between them.

3.1.1 Participants

18 participants (6 males, 12 females) were recruited for Experiment 1.1, all were native speakers of North American English and currently students at McGill University. Each participant received a compensation of C$ 12 per hour. All of them were naive as to the purpose of the experiment.

3.1.2 Material

In total there are 16 target sentences that each participant has to rate. These target sentences are controlled for two variables *extraction* (i.e. whether there is wh-extraction or not) and *fronting* (i.e. whether there is a fronted DP within the embedded clause or not), thus providing a total of four different conditions. The stimuli were not controlled for length of a condition across different items. Example (23) illustrates one such item set. All stimuli in this sub-experiment are presented out of the blue, i.e. without a prior context.

(23)   a.   I think he attended the **wedding** of the king.
          [CONDITION 1: NO FRONTING, NO EXTRACTION]
    b.   Which royalty do you think he attended the **wedding** of?
          [CONDITION 2: NO FRONTING, EXTRACTION]
    c.   I think the **wedding** of the king, he attended.
          [CONDITION 3: FRONTING, NO EXTRACTION]
    d.   Which royalty do you think the **wedding** of, he attended?
          [CONDITION 4: FRONTING, EXTRACTION]

3.1.3 Prosodic labeling

All intonation contours on target sentences produced by participants were labeled either as having default, shifted, double focus, no focus or other intonation patterns. Each labeling will be explained one by one. The *no focus* labeling, as the name already suggests, means that there was no prosodically focused element in the sentence perceived.

A default prosody places the prominence on the final lexical head, see Example (24) (the focused constituent has been written in boldfaced):

(24)   *Default prosody*
     a.   I think he attended the **WEDDING** of the king.
     b.   Which royalty do you think he attended the **WEDDING** of?
     c.   I think the wedding of the king, he **ATTENDED**?
     d.   Which royalty do you think the wedding of, he **ATTENDED**?

The shifted prosody is in some sense the opposite of the default prosody, the prominence of the sentence is placed on the second to last head, see (25).

(25)   *Shifted prosody*
     a.   I think he **ATTENDED** the wedding of the king.
     b.   Which royalty do you think he **ATTENDED** the wedding of?
     c.   I think the **WEDDING** of the king, he attended.
     d.   Which royalty do you think the **WEDDING** of, he attended?

The double focus prosody combines both default and shifted prosody and places prominence on both the final head and second to last head (26). Lastly, any other kind of contour pattern is labelled as *other*.

(26)   *Double focus prosody*
     a.   I think he **ATTENDED** the **WEDDING** of the king.
     b.   Which royalty do you think he **ATTENDED** the **WEDDING** of?
     c.   I think the **WEDDING** of the king, he **ATTENDED**.
     d.   Which royalty do you think the **WEDDING** of, he **ATTENDED**?

3.1.4 Procedure

The experiment was conducted in the Prosody Lab at McGill University and was created using the Psychtoolbox in Matlab. The participants were seated in a sound-attenuated booth in front of a computer. They received written instructions asking them to first carefully read the sentences, then produce them as natural as possible and judge the sentences on a eight-point scale (1 being completely natural and 8 being completely unnatural).[5] In a sense the task is very similar to a regular acceptability judgment task, except that besides just reading the sentences participants were asked to produce them.

By producing the sentences themselves, the participants presumably become more aware about how easy or hard it is to produce these sentences which will hopefully be reflected in the acceptability judgments. The reason for not simply asking the participants to read the sentences carefully was the fear that participants would potentially read over such a small word order difference like DP-fronting. Having to produce the sentences besides just reading them enforces the participants to read the sentences more carefully.

After reading the instructions, the first trial started and the participants were visually presented with the target sentence to be produced. After carefully reading the sentences, participants were asked to press the space key whenever they wanted to start recording the sentence. No correction of the recording could be made. After 5000 ms the input screen appeared automatically in which the participants were asked to judge the sentence. By pressing any key, the next trial would start. Participants proceeded through the experiment at their own pace. The participants were under no time pressure as to ensure they had enough time to process the

---

5   As a safeguard against floor effects, the choice was made to use an eight-point scale rather than the more standard seven-point scale.

sentences. The whole experiment took about 10 minutes.[6] Note that participants did not get to hear their own productions, the production data was collected purely for analysis reasons.

The stimuli are presented to the participants in a Latin square design meaning that no participant saw the exact same item twice. Each condition is repeated four times and each repetition is taken from a different item set, making a total of 16 observations per participant.

3.1.5 Results

The results from Experiment 1.1 are illustrated in *Figure 2*. A total of 288 observations were collected, i.e. 18 participants*16 observations = 288 observations. As becomes clear from this figure, the *No Fronting* condition seems more acceptable than the *Fronting* condition: CONDITION 1 (M = 6.94, SD = 1.35, 95% CI [6.63, 7.26]), CONDITION 2 (M = 5.40, SD = 2.10, 95% CI [4.91, 5.90]), CONDITION 3 (M = 3.94, SD = 2.09, 95% CI [3.45, 4.44]), CONDITION 4 (M = 2.99, SD = 1.79, 95% CI [3.41, 2.56]). In addition, there is a smaller effect of wh-extraction: Having wh-extraction is less acceptable than having no wh-extraction.
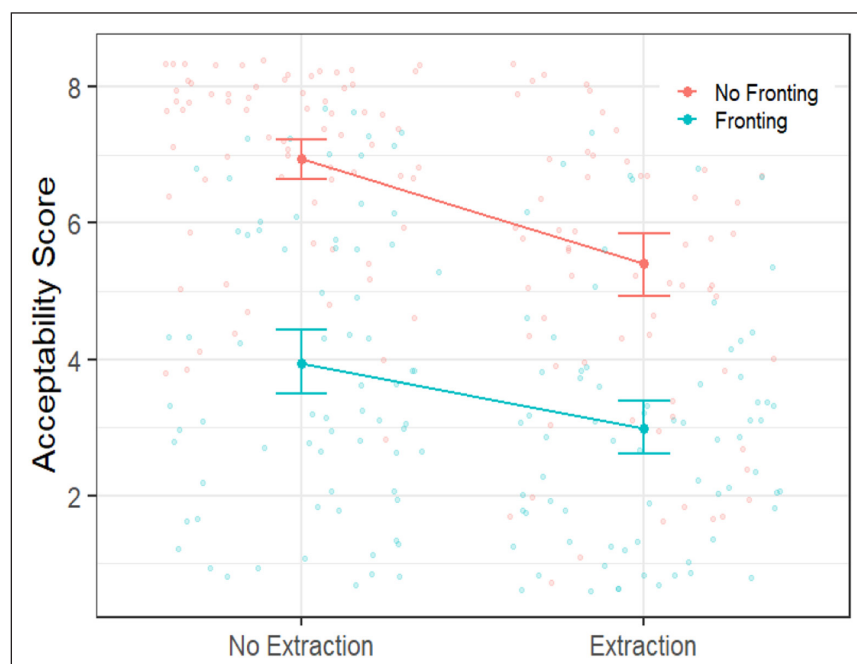
Figure 2 Results Experiment 1.1. Error bars show the 95% confidence intervals.

To see whether the observations from the figure are significant a linear mixed effects analysis was run using R (R Core Team 2018) and the LMER function in the lme4 Rpackage (Bates et al. 2014). As fixed effects *fronting, extraction* and their interaction were added to the model. As random effects, by-subject and by-item random intercepts and slopes for all fixed main effects were included in the model (See *Appendix B: Statistical Models* to see the actual model). Both *fronting* and *extraction* improved the model in comparison to a model in which these variables were excluded. The full model without interaction effects (AIC 1140) significantly improved the model compared to a model without fronting (AIC 1171.1) (χ^2 (1, N = 288) = 33.16, $p <$ 0.001)*:* Sentences with fronted DPs are significantly less acceptable than non-fronted ones ($b =$ –3.0, se = 0.344, t(35.96) = –8.725, p = 2.10e–10).

Similarly, the full model was a significant improvement to a model without extraction (AIC 1162) (χ^2 (1, N = 288) = 24.06, p < 0.001): Sentences with extraction are less acceptable than sentences without extraction (b = –1.547, se = 0.27, t(61.28) = –5.743, p = 3.11e–07).

The full model with interaction effects (AIC 1139.4) did not significantly improve the model compared to a model without an interaction effect between fronting and extraction (AIC 1140) (χ^2 (1, N = 288) = 2.55, p = 0.1103). Crucially, no significant interaction effect between fronting and extraction was found.

---

6   This experiment was performed as part of a big pool of several unrelated experiments. So even though the participants only took 10 minutes to finish this experiment, they spend a total of one hour each performing several experiments consecutively.

3.1.6 Prosodic results

Figure 6 shows a proportion graph of the productions made by the participants. A total of 263 observations were collected.[7]

A quick look at the prosodic data shows that different participants apply different strategies to place prominence on a certain element, some apply higher intensity (e.g. *Figure 4*) while others apply a higher pitch (e.g. *Figure 3*). Still others do not place any focus in the sentence and produce a rather mono-tonal utterance (see *Figure 5*).[8] These have been labeled as *other intonation* and form only a small proportion of the produced intonations as we can see from *Figure 6*. Due to the low number of non-default intonations no further analyses were performed on the production data.
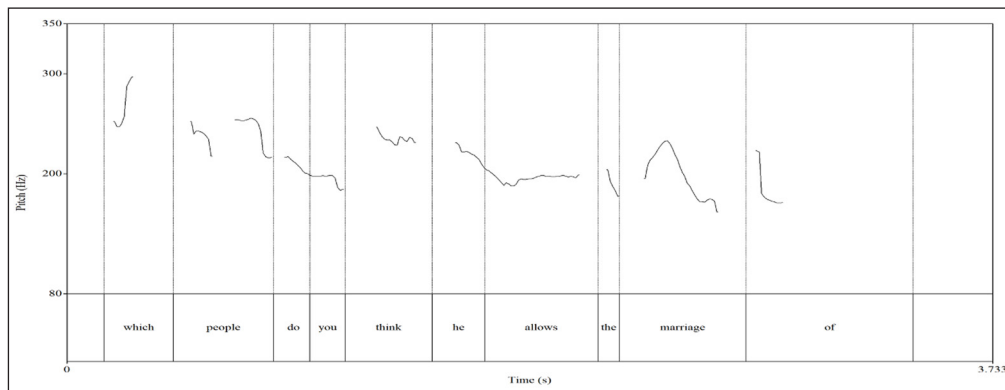
Figure 3 Production data without Fronting from female participant 9185. The line drawn indicates the pitch track. From the pitch track it becomes clear that the noun *marriage* receives main accent.
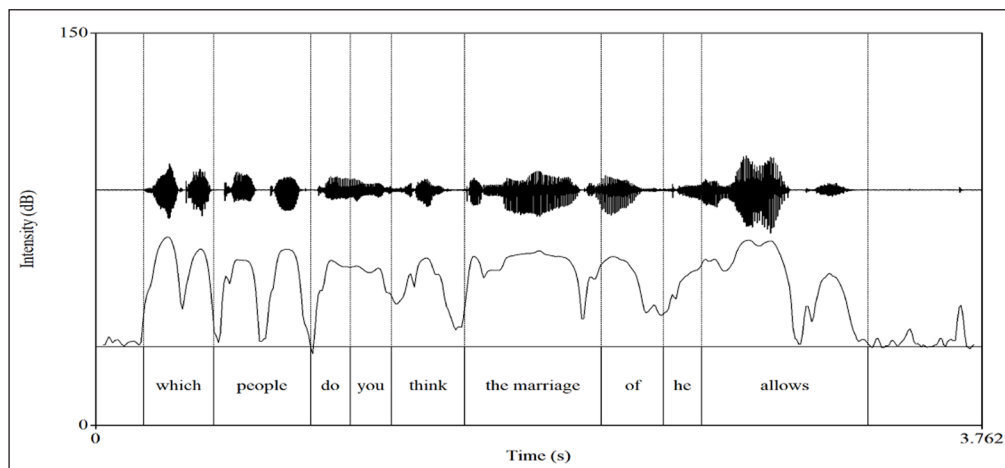
Figure 4 Production data with Fronting from female participant 9171. The line drawn underneath the waveform indicates the intensity. Both from the waveform as well as the intensity drawing it becomes clear that the verb *allows* receives main accent.
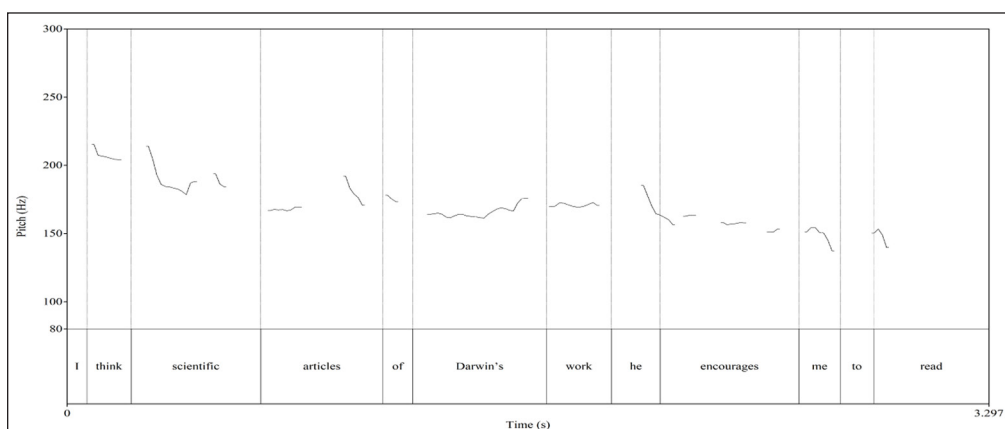
Figure 5 Example production by female participant 1970. The line drawn indicates the pitch track. It becomes clear that the sentence was produced in a very monotonous manner.

---

7   25 observations had to be removed due to the audio file either being cut off or the participants making background noises (like tapping on the table or making a lot of head movements) which interrupted the audio signal.

8   As noted by one reviewer a monotonous intonation may suggest shallow processing by the participant. This is definitely an issue, luckily not many monotonous intonations were produced.
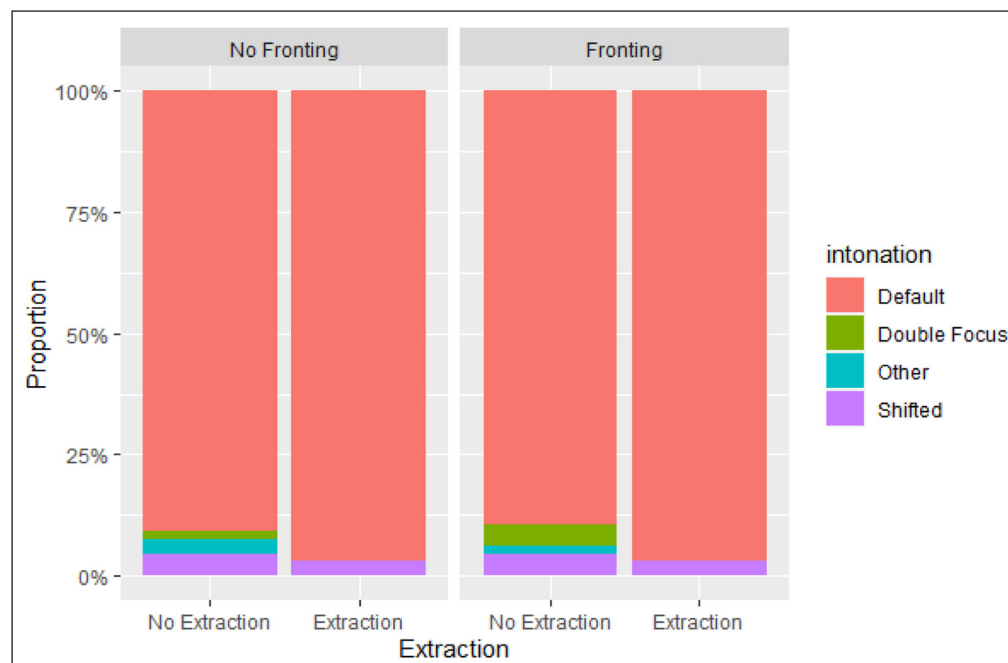
3.1.7 Conclusion experiment 1.1

The results from experiment 1.1 show a significant main effect of fronting and extraction but no interaction effect between the two variables. All these results support an additive account of the freezing effect. If freezing was truly a phenomenon on its own, one would have expected an interaction effect between fronting and extraction which would indicate that the lower acceptability of frozen sentences cannot simply be predicted from the individual contributions of fronting and extraction on the acceptability judgments. More specifically, one would have expected the difference in fronting to be larger in the *Extraction* condition than in the *No-extraction* condition, i.e. a super-additive effect. Instead of a super-additive effect, evidence for an additive effect was found which seems to indicate that the freezing effect is not a phenomenon on its own but a combination of two processes lowering the acceptability judgments (i.e. wh-extraction and fronting). The fact that there was no interaction effect does not automatically rule out a super-additive effect, however as becomes clear from *Figure 2*. There is an interaction trend in the opposite direction (i.e. the combination of fronting and extraction is less bad than would have been predicted based on the combination of the two effects) which makes a super-additive effect less likely.

Hofmeister et al. (2015) and Konietzko et al. (2018) conducted an experiment similar to experiment 1.1 to test for additivity and took the results to be evidence for a processing account of freezing. Contrary to those studies, the current study does not take additivity to be conclusive evidence for a processing account of freezing. Instead, experiment 1.1 is taken as a baseline to explore which factors contribute to the freezing effect. More specifically, two factors will be looked into namely prosodic and pragmatic factors. According to a prosodic account of freezing, extraction is only allowed from a focused constituent. Without a contrastive context, one expects that the sentences are produced using a default intonation pattern.

As becomes clear from the proportion graph for experiment 1, this is exactly the case (see *Figure 6*. With a default intonation the main accent falls on the DP in the *No Fronting* condition (see *Figure 3* and on the main verb in the *Fronting* condition (see *Figure 4*. This means that the main prominence of the sentence in the fronted condition would be on the main verb rather than the moved DP (see Examples (24d) and (24c). This could potentially explain why the fronted condition is significantly less acceptable than the non-fronted condition. The prosodic data thus seem to support a prosodic account, however as will become clear later this is not the only explanation.

In line with previous experimental research on freezing, Experiment 1.1 supports the idea that freezing is best described as an additive effect. In addition, the production data are compatible with a prosodic account. Different from the prosodic account, I propose that lower acceptability

ratings are caused by pragmatic factors rather than prosodic ones. Experiment 1.2 will test the prosodic account and compare it against a pragmatic account. In addition, Experiment 1.2 will tell how much of the additive effect can be ascribed to the presence of a context: Do both DP-fronting and wh-extraction improve by context, or just one of the two?

According to previous literature on DP-fronting, one of the requirements of fronting is the need for a contextually given referents (Ward & Birner 1994; Birner & Ward 1998). Since there was no such referent in Experiment 1.1, nor in the previous studies by Hofmeister et al. (2015) and Konietzko et al. (2018), the lower acceptability ratings for the *Fronting* condition might as well be ascribed to the lack of such a contextual anchor.

Similarly, one could imagine that certain wh-question require a contextually given referents. All the stimuli in Experiment 1.1 contain a wh-word which presuppose the existence of a certain set of entities, or a so-called D-linked wh-word. *Which royalty* in Example 2 for instance, presupposes the existence of a specific royalty and the question might be less acceptable when no such royalty was mentioned in the previous context. To test for this, Experiment 1.2 was set up in which a context containing a referent was provided for each target sentence.

3.2 Experiment 1.2

In Experiment 1.2 a contrastive context is added to the target sentences in order to see the effect of context on freezing. As already mentioned in the introduction, there is a intuition that these freezing sentences become more acceptable with the right contrastive focus context. The explanation for why freezing improves in such cases has been related to the hypothesis that extraction can only take place from a focused constituent (Bayer 2005; Culicover & Winkler 2010; Winkler et al. 2016). Experiment 1.2 sets out to test this hypothesis by using the same target sentences from Experiment 1.1 in two different contrastive focus contexts, i.e. contrastive DP focus and contrastive VP focus. In the contrastive DP context, a DP appears in the preceding context which contrasts with the potentially fronted DP in the target sentence. This context is expected to improve DP-fronting. In the contrastive VP context, a verb in the preceding context contrasts with a verb in the target sentence: No improvement of DP-fronting is expected under the prosodic account.

More specifically, the prosodic account will be tested against a pragmatic account. The pragmatic account holds that DP-fronting requires a contextually given referents and the absence of such referent accounts for the apparent freezing effect. We will see that all contexts in Experiment 1.2 contain a contextual anchor with which the DP can associate. If the prosodic account is on the right track, one expects a significant effect of the context condition according to which only the DP-context significantly improves fronting, but not the VP-context. If on the other hand the pragmatic account is correct, no such significant effect should arise. In addition, this should be reflected in the prosody as well. If the prosodic account is on the right track, one expects DP-fronting to be more acceptable when that DP has been accented. For the pragmatic explanation to be true, no difference between the specific focus conditions is expected, but there should be a significant difference between Experiment 1.1 and Experiment 1.2, i.e. sentences without a contextually given referents versus sentences with a contextually given referents. Following the pragmatic account we expect to find a significant effect of context when comparing Experiment 1.1 with 1.2, but no significant effect between the two context conditions as would have been predicted by the prosodic account.

3.2.1 Participants

36 participants (16 males, 20 females) were recruited for experiment 1.2, all were current students at McGill University. Each participant received a compensation of C 12 per hour. The group of participants was split up in two groups, one group received target sentences with wh-extraction and the other one without wh-extraction. The reason for doing so was to keep the duration similar to experiment 1.1. 14 of them participated in the first sub-experiment and 12 of them participated in the second sub-experiment. All participants were naive as to the purpose of the experiment.

3.2.2 Materials

16 different item sets were created for this experiment. One such item set is given below in Example (28) and (27).[9] In total there are 16 target sentences that the participant is presented with and has to rate. These target sentences are the same as the once in experiment 1.1, this way the syntax is kept constant across the two experiments.

Each item set is defined for two variables with each two different settings, these are focus (i.e. *DPFocus* and *VPFocus* and fronting (i.e. fronted or non-fronted). To make sure the correct parts of the sentences are focused across the different conditions, each target sentence is preceded by a pre-sentence which will lead the listener to the intended interpretation of the sentence.

In order for the pre-sentences to have a contrastive effect on the target sentences, two contrastive item pairs are introduced, one contrastive DP pair and one contrastive verb pair. In this case for the DP pair, *engagements* is contrasted with *weddings* and for the contrastive verb pair *attend* is contrasted with *missed.* The other variable that has been manipulated in the stimuli set is fronting. In Example (28a), there is a contrastive focus on *wedding* which is contrasted with *engagement* in the pre-sentence. Example (28c) is similar to Example (28a) except that the DP *the wedding of* has been fronted within the embedded clause. The contextual anchor to the fronted DP *wedding* is in this case the contrasted item *engagement.*

Moving on to contrastive verb focus, in Examples (28b) and (28d) the verbs *missed* and *attended* are contrasted with each other. Even though the verbs in this case are contrastive rather than the DPs, there is still a contrastive contextual anchor to the fronted DP. The contrast here is the wedding that was attended versus the wedding that was missed. The contextual anchor to the fronted DP *(attended) wedding* is in this case the contrasted item *(missed) wedding* in the pre-sentence. This set up helps us distinguish prosodic factors from pragmatic ones as we will see later on in the paper. The sentences in Example (27) are similar to the ones in (28) with the exception that there is no wh-extraction. It should also be noted that each stimulus is preceded by a context which introduces the topic of the conversation. The prosodic data is labeled similarly as in Experiment 1.1.

Context: *Being a prince himself, George is often invited for engagements and weddings of other royalties in his family. However, he is too busy to attend all engagements and weddings, so he missed a few weddings and engagements.*

(27)   *No Extraction*
   a.   It's not that George attended the **engagement** of the king,…I think he attended the **wedding** of the king.
           [CONDITION 1: DPFOCUS, NO FRONTING, NO EXTRACTION]
   b.   It's not that George **missed** the wedding of the king,…I think he **attended** the wedding of the king.
           [CONDITION 2: VPFOCUS, NO FRONTING, NO EXTRACTION]
   c.   It's not that George attended the **engagement** of the king,…I think the **wedding** of the king, he attended.
           [CONDITION 3: DPFOCUS, FRONTING, NO EXTRACTION]
   d.   It's not that George **missed** the wedding of the king,…I think the wedding of the king, he **attended**.
           [CONDITION 4: VPFOCUS, FRONTING, NO EXTRACTION]

(28)   *Extraction*
   a.   I know which royalty George attended the **engagement** of, but…which royalty do you think he attended the **wedding** of?
           [CONDITION 5: DPFOCUS, NO FRONTING, EXTRACTION]
   b.   I know which royalty George **missed** the wedding of, but…which royalty do you think he **attended** the wedding of?
           [CONDITION 6: VPFOCUS, NO FRONTING, EXTRACTION]
   c.   I know which royalty George attended the **engagement** of, but….which royalty do you think the **wedding** of, he attended?
           [CONDITION 7: DPFOCUS, FRONTING, EXTRACTION]

---

9   See Appendix B for a full list of the stimuli used in this experiment as well as the instructions.

d. I know which royalty George **missed** the wedding of, but…which royalty do you think the wedding of, he **attended**?
[Condition 8: vpfocus, fronting, extraction]

3.2.3 Procedure

The experiment is set up according to a between-subject design which means that one group of participants gets to see the extracted sentences and the other group the non-extracted ones. The reason for doing so is to make sure that there was no effect of trial due to boredom and to keep it as long as experiment 1.1. In the instructions the participants were told they would see 16 sentences which needed to be produced and then rated on a scale from 1 to 8 (1 being completely unnatural and 8 being completely natural). The stimuli are presented to the participant in a Latin square design, i.e. all participants see each condition. Each condition is repeated four times, but since they are picked from different item sets, there will be no identical trials. They were told that each sentence is preceded by a context. After reading the instructions, the participants were auditorily presented with a context. After the context ended, the experiment automatically continued to the next screen. The next screen showed the pre-sentence and target sentence to be produced separated by three dots. After carefully reading the whole sentence, participants were asked to press the space key to start recording the whole sentence, i.e. pre-sentence and target sentence (to make the recording more natural). After 5000 ms the input screen appeared automatically in which the participants were asked to judge only the last part of the sentence, i.e. the part following the three dots.[10] By pressing any key, the next trial would start. Participants proceeded through the experiment at their own pace. The participants were under no time pressure as to ensure they had enough time to process the sentences. The whole experiment took about 10 minutes.

3.2.4 Results

A total of 576 observations were collected, i.e. 36 participants * 16 stimuli = 576 observations. As becomes clear from *Figure 7*, there seems to be a clear interaction effect between fronting and extraction according to which the difference in fronting is larger in the *Extraction* condition than in the *No Extraction* condition. The focus manipulation on the other hand, does not seem to have a clear effect on the data: condition 1 (M = 5.47, SD = 1.98, 95% CI [5.01, 5.94]), condition 2 (M = 4.66, SD = 2.35, 95% CI [4.11, 5.21]), condition 3 (M = 4.35, SD = 2.15, 95% CI [3.84, 4.85]), condition 4 (M = 4.34, SD = 2.10, 95% CI [3.85, 4.84]), condition 5 (M = 5.81, SD = 1.72, 95% CI [5.40, 6.21]), condition 6 (M = 5.75, SD = 1.68, 95% CI [5.35, 6.14]), condition 7 (M = 3.86, SD = 1.82, 95% CI [3.43, 4.29]), condition 8 (M = 4.00, SD = 2.02, 95% CI [3.52, 4.48]).
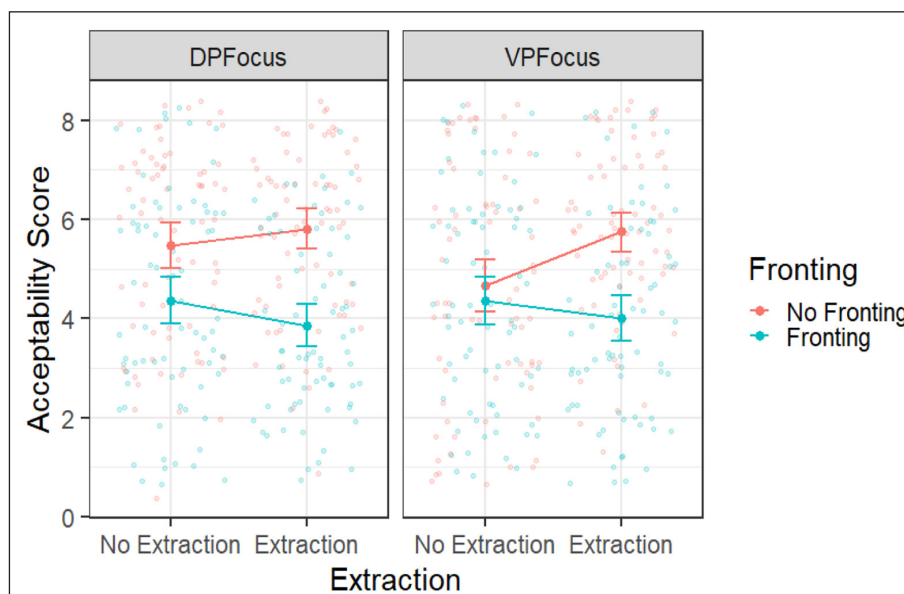
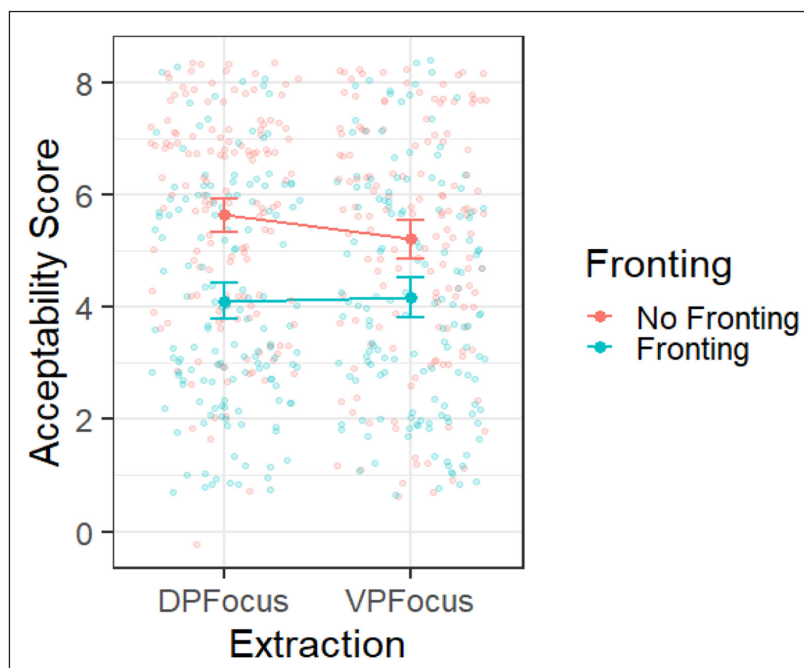**Figure 7** Results experiment 1.2. Error bars show the 95% confidence intervals.

---

10 This is to ensure they are rating the same sentences as the ones in Experiment 1.1.

To see whether the observations made from the figure are significant a mixed linear effects analysis was run. In the final model focus, extraction, fronting and the interactions between fronting and focus, and extraction and fronting were added to the model. As random effects, by-subject and by-item random intercepts and slopes for all main effects were included in the model.[11] All models were compared to the full model which had as fixed effects focus, extraction, fronting and the interaction of all these variables, and ss random effects, by-subject and by-item random intercepts and slopes for all main effects. The model comparison shows that only fronting improved the model in comparison to a model in which these variables were excluded. The full model without interactions (AIC 2328.6) was a significant improvement to a model without fronting (AIC 2354.7) (χ^2 (1, N = 576) = 28.05, p = 1.179e–07): The *Fronting* condition is significantly less acceptable than the *Fronting* condition (b = –0.985, t(59.06) = –3.4, P = 0.0012).

The full model without interactions (AIC 2328.6) did not perform better than the model without focus (AIC 2327.6) (χ^2 (1, N = 576) = 0.972, p = 0.3241), nor did the full model without interactions improve a model without extraction (AIC 2327.6) (χ^2 (1, N = 576) = 1.021, p = 0.3124).

Different from Experiment 1.1 though, the full model with interactions (AIC 2320.1) performed significantly better than the model without an interaction effect between extraction and fronting (AIC 2327) (χ^2 (1, N = 576) = 8.903, p = 2.85e–3): This significant result indicates that the difference in fronting is larger in the *Extraction* condition than in the *No Extraction* condition (b = –1.125, t(37.61) = –3.084, p = 0.004).

Although no significant main effect of focus was found there was an interaction effect between focus and fronting. The full model performed significantly better than a model without this interaction effect (AIC 2322.1) (χ^2 (1, N = 576) = 3.964, p = 0.046): The difference in fronting is bigger in the DPFocus condition than the VPFocus condition, as can be seen more clearly in *Figure 8* (b = 0.525, t(459.71) = 1.99, p = 0.0473).

Figure 8 Results experiment 1.2 without extraction. Error bars show the 95% confidence intervals.

3.2.5 Prosodic results

In contrast to experiment 1.1 where most sentences were produced with a default intonation pattern, Experiment 1.2 shows a clear prosodic pattern depending on the different conditions (see *Figure 9*. A total of 488 observations were collected.[12]

---

11   See *Appendix B: Statistical Models* to see the actual model.

12   This means that a total of 88 (576 total observations-488 acceptable observations = 88) observations have been excluded, for similar reasons as in Experiment 1.1.
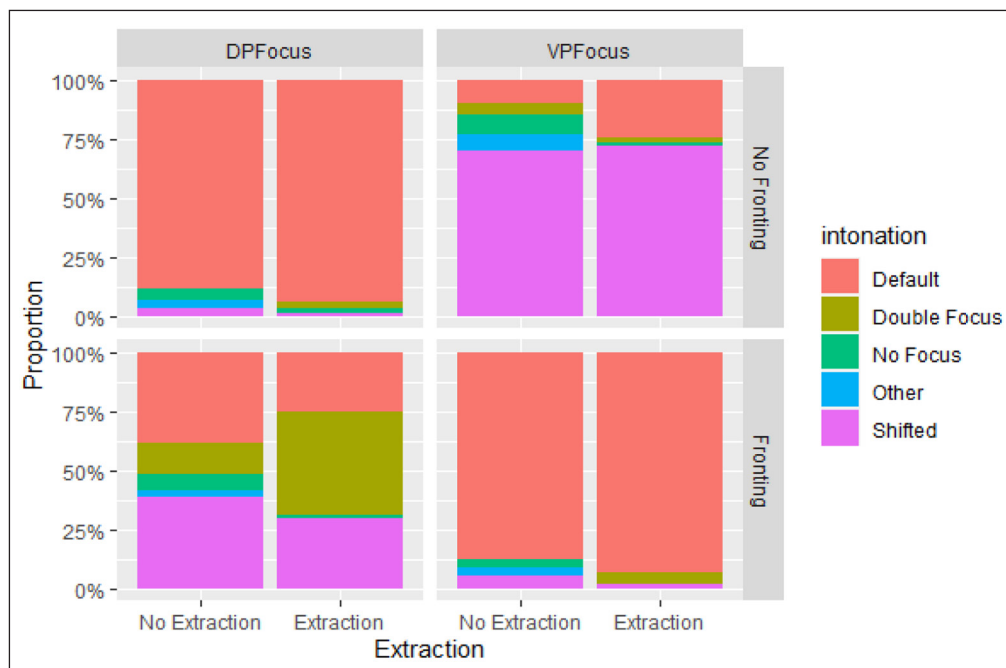
To see whether intonation had a significant effect on the acceptability ratings a mixed linear effects analysis was run. For the fixed effects, focus, extraction, fronting and intonation were added to the model as well as the interaction between fronting and extraction. As random effects, by-subject and by-item random intercepts and slopes for all main effects were included in the model.[13] The model comparison shows that intonation did not improve the model nor did any interactions between intonation and the other variables. The full model without interactions (AIC 1918.1) was not a significant improvement to a model without intonation (AIC 1916.6) (χ^2 (1, N = 488) = 0.5379, p = 0.4633). In addition none of the interactions with intonation significantly improved the model: Adding the interaction between intonation and fronting to the model (AIC 1915.8) did not significantly improve the full model (AIC 1913.9) (χ^2 (1, N = 488) = 0.0937, p = 0.7595). The full model also did not significantly improve when adding the interaction between intonation and extraction to the model (AIC 1915.8) (χ^2 (1, N = 488) = 0.0486, p = 0.8256). Nor did a three-way interaction between intonation, fronting and extraction (AIC 1914.3) significantly improve the full model model (χ^2 (1, N = 488) = 1.5593, p = 0.2118).

3.2.6 Conclusion experiment 1.2

So far we have seen that for Experiment 1.2 there is a significant effect of fronting on the acceptability of the sentence: sentences with fronted DPs are less acceptable than the ones without fronting. Interestingly, there is an interaction effect between fronting and extraction. This interaction effect indicates that the difference in fronting is larger in the *Extraction* condition than in the *No Extraction* condition. What is unexpected though is the direction of this interaction effect: *Fronting* seems to become less acceptable with wh-extraction (as would have been expected), but the *No Fronting* condition becomes better with wh-extraction. So far, I have no reasonable explanation for this. More interestingly, the results tell us that there is a significant interaction effect between the focus manipulation and fronting. This is exactly what the prosodic account would have predicted, however, the interaction does not have the expected effects. According to the prosodic account the *DPFocus* contexts should enhance fronting, whereas the *VPFocus* condition would not: There is no difference in fronted sentences between the *DPFocus* and *VPFocus*. What one observes instead is that the *No Fronting* condition becomes more acceptable with a *DPFocus* context, see *Figure 8*. It is unclear why this would be the case, one possible explanation could be that by default the accent in a non-fronted sentence is placed on the DP. The VPFocus would then clash with the default intonation since it would place stress the accent on the verb instead. Crucially, there is an effect of focus on the acceptability of the sentence, however this only affects the non-fronted sentences but not the

---

fronted sentences. This all seems to suggest that one is not dealing with a prosodic garden path as advocated by Culicover & Winkler (2010).

The prosodic data provide even more evidence against a prosodic account. The expectation is that the contrastive contexts should force the speaker to prosodically focus either the DP or verb with which a contrast is made. As expected, the *No Fronting* condition with a *DPFocus* context has a high percentage of default intonation, whereas the ones with *VPFocus* contexts have a higher proportion of shifted intonation. This means that in the *DPFocus* condition the main accent is placed on the contrasted DP and in the *VPFocus* condition the main accent is placed on the contrasted verb. For the Fronting condition one would expect the exact opposite pattern. In a fronted sentence one observes a high percentage of default intonation with a *VPFocus* context. With default intonation the main accent would fall on the main verb. Less expected are the fronted sentences with a *DPFocus* context. One would have expected to see considerably more shifted intonation than default intonation, since that would place main accent on the contrasted DP. This is not the case and there is no good explanation yet for this. However what is most important is that there is a difference between the two focus conditions in production, but not in acceptability scores. In other words, even though in most cases the expected intonation pattern is produced and a difference is made between the *DPFocus* and *VPFocus* conditions, no such difference is reflected in the acceptability data.

The results from Experiment 1.2 are compatible with a pragmatic approach. If the pragmatic account is on the right track one expects to find an improvement of adding a context for the *Fronting* condition but not for the *No Fronting* condition. Something similar could also be expected for wh-extraction, if its really true that wh-extraction requires a prior context.

3.3 Context vs No-Context

Even though from experiment 1.2 it is concluded that the focus manipulation (i.e. either VP focus or DP focus) did not have the expected effect on the acceptability ratings of the sentence, we will see that there is an effect of context in general which makes more sense. To see whether there is an effect of context on the acceptability ratings, experiments 1.1 (without context) and 1.2 (with context) were compared with each other. The results are shown in *Figure 10* below. A total of 864 observations have been collected, i.e. observations of experiment 1.1 + observations of experiment 1.2.
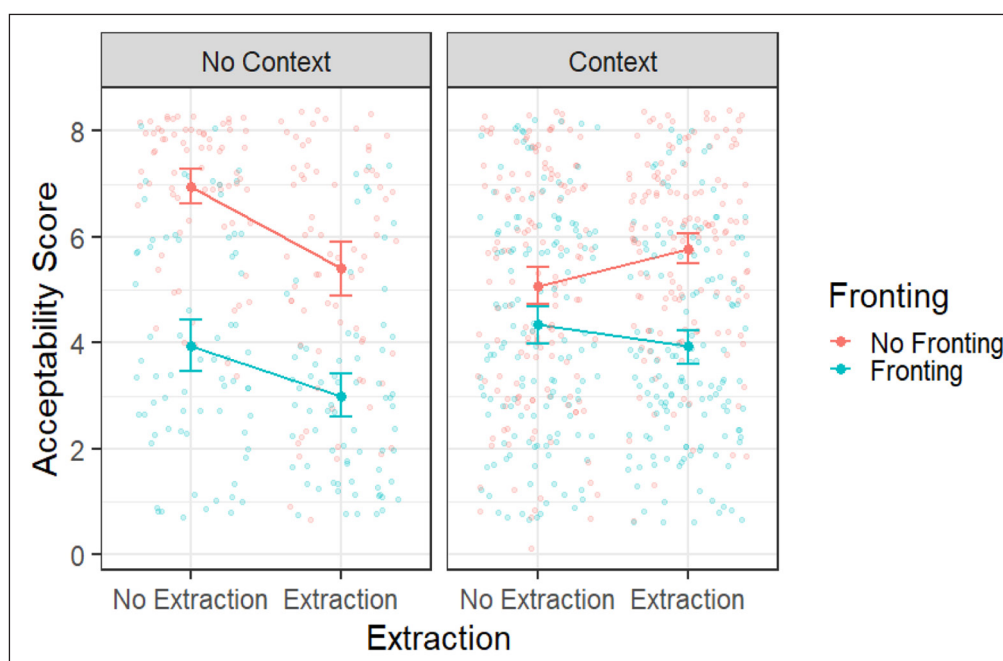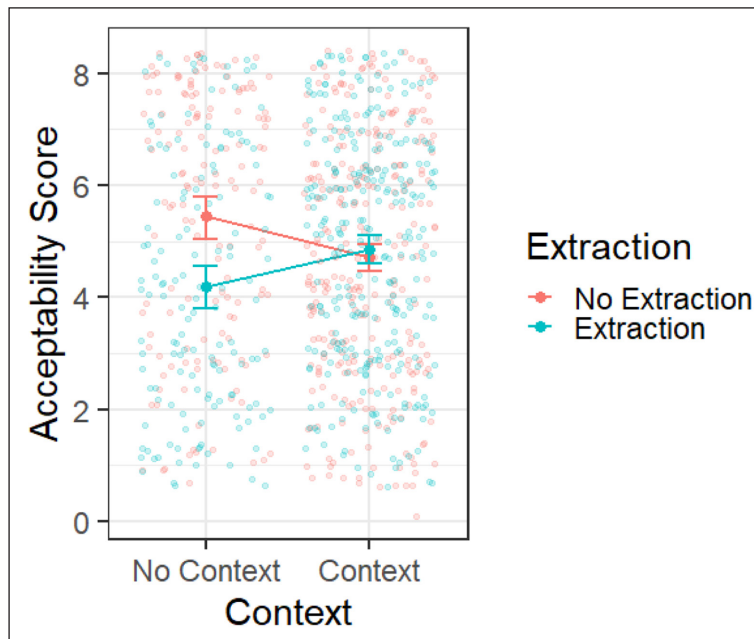

Figure 10 Results experiment 1.1 and 1.2. Error bars show the 95% confidence intervals.

Once again a mixed effects model was run: For the fixed effects, extraction, Fronting, presence of context and the interactions between the different variables were added to the model as well as a threeway interaction between all three main effects. As random effects, by-subject and by-item random intercepts and slopes for all main effects were included in the model.

Crucially, the mixed linear effects model shows that there is a significant three-way interaction between fronting, extraction and context. The full model (AIC 3451.4) performed significantly better than a model without this three-way interaction (AIC 3458.5) (χ^2 (1, N = 864) = 9.095, p = 0.0025): Context significantly changes the effect that fronting has on extraction (b = –1.6198, t(223.18) = –3.017, p = 2.85e–3).

In addition, there is a significant interaction effect between context and extraction. The full model performed significantly better than a model without this interaction effect (AIC 3472.2) (χ^2 (1, N = 864) = 22.814, p = 1.784e–06): This result indicates that the difference in extraction is bigger when there is no context as opposed to when there is a context (b = 2.2819, t(75.168) = 5.102, p = 2.44e–06), see *Figure 11*.

Figure 11 Results experiment 1.1 and 1.2 without fronting. Error bars show the 95% confidence intervals.

Similarly, there is a significant interaction effect between *fronting* and *context*. The full model (AIC 3451.4) performed significantly better than a model without this interaction effect (AIC 3474.3) (χ^2 (1, N = 864) = 24.95, p = 5.892e–07): This result indicates that the difference in fronting is bigger when there is no context as opposed to when there is a context (b = 2.192, t(126.23) = 5.27, p = 5.72e–07), see *Figure 12*.
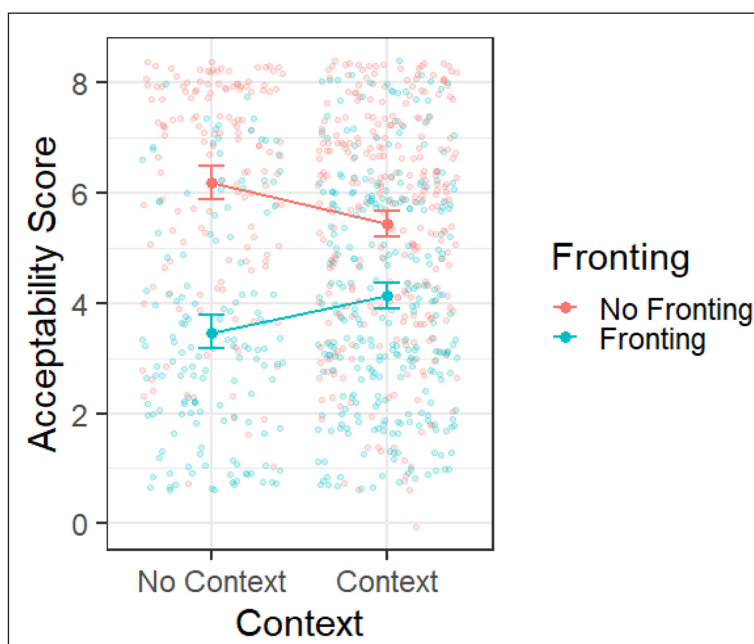
Figure 12 Results experiment 1.1 and 1.2 without extraction. Error bars show the 95% confidence intervals.

3.4 Conclusion experiment 1

In this experiment we have seen that the effects of wh-extraction and DP-fronting can be added up to account for the freezing effect (Experiment 1.1). Moreover, it has become clear that contrastive DP focus does not improve frozen sentences more than a contrastive verb focus (in fact the opposite was observed), as would have been predicted according to the prosodic account (Experiment 1.2). Context in general, on the other hand, does have an effect on frozen sentences (comparison between experiment 1.1 and 1.2): Adding a context improves the *Fronting* condition, but it also lowers the *Nofronting* condition. The fact that context improves fronting is expected, however it is less clear why adding a context would make the *No Fronting* condition less acceptable. Potentially, the contexts are set up in such a way that they are biased towards a fronted construction, thus making the non-fronted constructions less suitable with such a context. The same can be said of the interaction between context and extraction. The data are hard to interpret and it is still unclear what the effect of context exactly is and why it affects the data the way it does. One conclusion that can be drawn though is that adding a context significantly changes the results. More importantly, both the *Fronting* condition as well as the *Extraction* condition seem to improve with context. Since the target sentences in both experiment were the same and therefore their syntax is the same, we can conclude that this improvement should be attributed to adding context rather than attribute it to structural differences.

A similar conclusion was drawn by Winkler et al. (2016) who also pointed out that depending on the context, freezing sentences seem to improve. Differently from their results though, no effect of the specific type of contrastive focus was found in the actual production data. Winkler et al. (2016) compared sentences with a contrastive focus context on the potentially moved constituent to the ones without such context. In contrast to Winkler et al. (2016), the current study compared two different contrastive focus contexts: one in which the potentially moved DP was contrasted and one in which the main verb was contrasted. According to Winkler et al. (2016) one would have expected the *VPFocus* context to behave similarly to the experiment without a context. However, it turns out that frozen sentences also improve even when a different constituent other than the fronted DP is contrasted. Perhaps the reason that Winkler et al. (2016) saw an improvement in the frozen sentences was merely due to the fact that there was a context with a contextually given referents rather than the fact that the prosody had been explicitly changed. The comparison between experiment 1.1 and 1.2 thus points towards a pragmatic explanation rather than a prosodic one.

Notice that even though Winkler et al. (2016) argue for a prosodic account of freezing, they did not actually look at production data explicitly but rather inferred indirectly from the acceptability judgments that their hypothesis was right. Comparing the two sub-experiments with each other it becomes clear that the fronted condition improves with a contrastive context. The reason for the lower acceptability judgments for the *Fronting* condition in experiment 1.1 compared to experiment 1.2 cannot be explained in terms of a clash between the default intonation contour and the information structure conveyed by the syntax, as claimed by Winkler et al. (2016). There is no correlation between the production data and the acceptability judgments to begin with. One might have expected for instance that a frozen sentence would be significantly more acceptable with a shifted contour than with a default contour. However this is not the case: The production data tell us nothing about the acceptability judgments.

The fact that there is no correlation between the production data and the acceptability judgments is unexpected for the prosodic account. We have seen that participants do produce the expected intonation contours based on the contrastive context preceding it, but their productions did not influence the acceptability judgments: the frozen sentences did improve with a contrastive DP context in which the moved DP from which extraction took place was accented, but so did a contrastive VP context. In other words it did not matter whether extraction took place from a focused constituent or not, the only thing that seemed to matter was the presence of a context containing an referential anchor.

We saw that experiment 1.1 mainly involved default intonation whereas in experiment 1.2 the focus shifted to the contrasted item. However it should be noted that for the condition in which prosody was expected to play the most important role, namely sentences with wh-extraction and a fronted embedded DP (i.e. frozen sentences) in a *DPFocus* context, the proportion of

shifted intonation contour was not as much as one would have predicted based on a prosodic account.

Even though comparing the production data with the acceptability judgment data seems to provide evidence against a prosodic account, the results might have been obscured by the complexity of the task. Note that the participant is first asked to produce the sentence, and then judge that sentence they just produced. This is a rather unnatural and complicated setting. The reason for setting up the experiment like this was to extract production data from the participants. To make the task less complex and more natural experiment 2 was run. This experiment is very similar to Experiment 1.2 with the exception that the stimuli are presented auditorily rather than forcing the subjects to produce them themselves. This should make the task easier and clearer patterns in the results are expected.

4 Experiment 2

Experiment 2 is similar to experiment 1.2 but instead of asking the participant to produce the target stimuli, the stimuli are auditorily presented. The participant's task is to assess how acceptable the target sentence sounds to him/her. Rather than having to come up with the right prosody for the target sentences themselves, they will be provided with the 'expected' prosody.[14] Another difference between experiment 1.2 and Experiment 2 is that the participants in this experiment performed both the stimuli with wh-extraction as well as without wh-extraction just like in Experiment 1.1.

Once again we expect a significant effect of context when comparing Experiment 1.1 to the results in this experiment, but no significant effect between the different context conditions as would have been predicted by a prosodic account.

4.1 Participants

A total of 18 participants were recruited for experiment 2 (7 males, 11 females). All of them were native speakers of North American English and were currently students at McGill University. Each participant received a compensation of C$ 12 per hour.

4.2 Materials

The materials and procedure are similar to experiment 1.2 except that participants have to rate pre-recorded stimuli. Both the context as well as the pre-sentence and target sentence were recorded for analysis purposes only. The target sentence was produced with emphasis on the constituent with which it contrasted in the pre-sentence, see Examples (28) and (27) for the produced patterns. The speaker mainly used intensity to mark a prosodic accent as becomes clear from *Figure 13*. The stimuli were recorded using Praat (Boersma & Weenink 2002) in a sound-attenuated booth (16 bit, 44100 Hz) by a male native speaker of American English (27 years old).
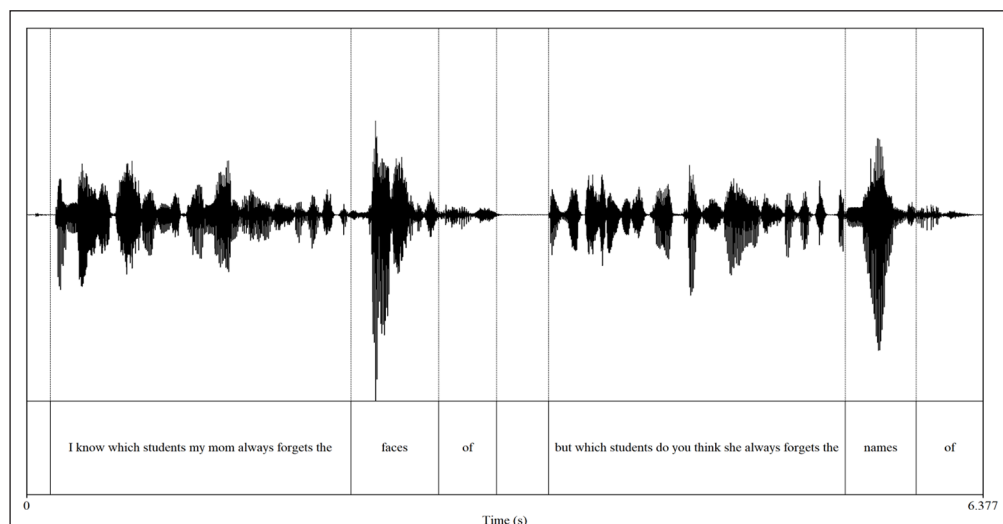
Figure 13 Example of a stimuli produced by a male speaker. The speaker uses intensity to mark a prominence. This becomes clear from the waveform.

14   Expected prosody here means that main stress is put on the contrastive lexical item in the sentence.

4.3 Procedure

The experiment was conducted in the Prosody Lab at McGill University and was created using the Psychtoolbox in Matlab. The stimuli were distributed according to a Latin square design meaning that no participant saw the exact same item twice. The participants were seated in a sound-attenuated booth in front of a computer. They received written instructions asking them to first carefully read the sentences, and judge the sentences after hearing them on a scale from 1 to 8. A total 32 stimuli were presented to the participant.

In the instructions the participants were told they would see 32 sentences which needed to be rated. After reading the instructions, the participants were presented with a context in written form. By pressing the space key, the next screen appeared which showed the pre-sentence and target sentence. After carefully reading the whole sentence, participants were asked to press the space key to hear the sentences. After 5000 ms the input screen appeared automatically in which the participants were asked to judge only the last part of the sentence, i.e. the part following the three dots. By pressing any key, the next trial would start. Participants proceeded through the experiment at their own pace. The participants were under no time pressure as to ensure they had enough time to process the sentences.

4.4 Results

A total of 576 observations were collected, i.e. 18 participants * 32 stimuli = 576 observations. Looking at *Figure 14* it becomes clear that fronting seems to have a significant effect on the acceptability of the sentences, i.e. sentences without fronting are more acceptable than sentences with fronting. Less obvious is the effect of extraction, it seems that there is no such effect, as becomes clear from the overlapping error bars: the pattern in the sentences with wh-extraction does not differ from sentences without extraction. Similarly, it is unclear what the effect of focus is on the acceptability rating: CONDITION 1 (M = 6.62, SD = 1.60, 95% CI [6.24, 7.00]), CONDITION 2 (M = 6.32, SD = 1.84, 95% CI [5.89, 6.75]), CONDITION 3 (M = 4.24, SD = 1.66, 95% CI [3.84, 4.63]), CONDITION 4 (M = 4.44, SD = 1.81, 95% CI [4.02, 4.87]), CONDITION 5 (M = 6.00, SD = 1.95, 95% CI [5.54, 6.46]), CONDITION 6 (M = 5.97, SD = 1.84, 95% CI [5.45, 6.40]), CONDITION 7 (M = 4.08, SD = 1.92, 95% CI [3.63, 4.53]), CONDITION 8 (M = 3.86, SD = 1.89, 95% CI [3.42, 4.31]).
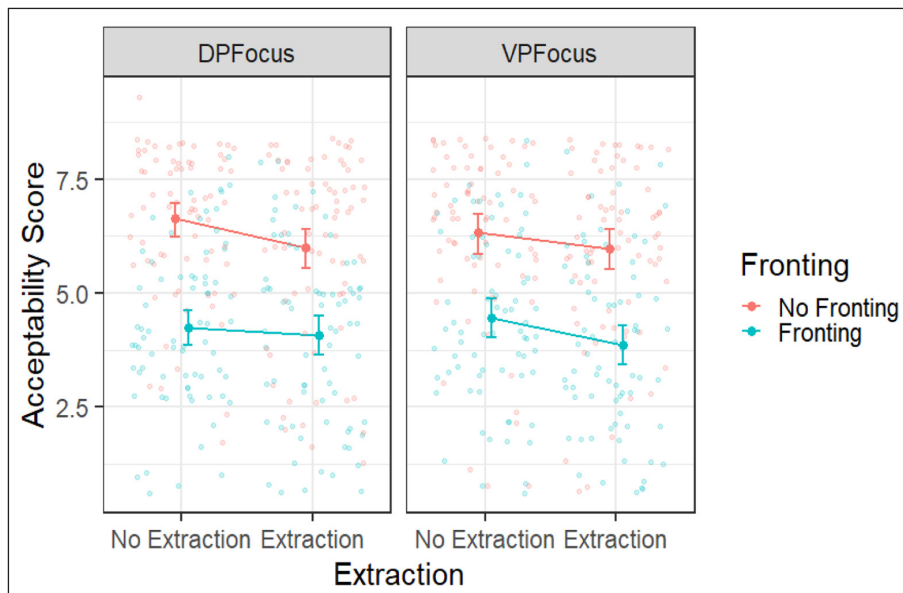
Figure 14 Results experiment 2. Error bars show the 95% confidence intervals.

To establish whether or not the observed differences are significant, a linear mixed effects analysis was performed using R (R Core Team 2018) and the LMER function in the lme4 Rpackage (Bates et al. 2014). For the mixed linear effects model, the variables focus, extraction and fronting were added as fixed effects in the model and their interactions. As random effects, by-subject and by-item random slopes for all main effects were included in the model. The model comparison shows that only fronting improved the model in comparison to a model in which these variables were excluded. The full model without interactions (AIC 2127.8) was

a significant improvement compared to a model without fronting (AIC 2151.8) (χ^2 (1, N = 576) = 26, p = 3.42e–07): Fronted sentences are less acceptable than non-fronted ones (b = –2.157, t(33.59) = –6.655, p = 1.3e–07). No improvement of the model was found for the other variables nor for any interactions between the variables.

There was no significant interaction effect between fronting and extraction. The full model (AIC 2132.8) was no improvement to a model without this interaction (AIC 2131.1) (χ^2 (1, N = 576) = 0.351, p = 0.55).

More importantly, there was no significant interaction effect between fronting and focus. The full model was no improvement to a model without this interaction (AIC 2131.4) (χ^2 (1, N = 576) = 0.57, p = 0.45).

4.5 Conclusion experiment 2

The results from experiment 2 tell us that sentences with fronted DPs are judged less acceptable than sentences without fronted DPs. This is a robust effect we have seen throughout all of the experiments.

One of the main reasons to perform Experiment 2 was to test the prosodic account. Different from Experiment 1.2 and Winkler et al. (2016), Experiment 2 directly tests the prosodic hypothesis by providing the participants with the hypothesized intonation contours. If the prosodic account is on the right track, then the different intonation patterns should have a significant effect on the acceptability judgments. This is not the case, there was no effect of focus on the acceptability judgments. Based on the prosodic garden path analysis of freezing sentences by Culicover & Winkler (2010) and Winkler et al. (2016), one would have expected that having the moved DP accented prosodically would have repaired a potential prosodic garden path and therefore fronted sentences with DP focus should be judged relatively more acceptable than fronted sentences with VP focus. The results of both Experiment 1 and Experiment 2 do not support the prosodic account.

The results are compatible with a pragmatic account, however, the fact that there was no effect of the specific contrastive focus context does not provide direct evidence for a pragmatic account. For a pragmatic account to be correct, it is expected that the *Fronting* condition becomes better by adding a context. The next section compares Experiment 1.1 and 2 to test this claim.

4.6 Context vs. No-context

Comparing Experiment 1.1 (No-context) with Experiment 2 (Context), gives us a clearer pattern than the previous comparison between Experiment 1.1 and 1.2, as can be seen from *Figure 15*. Whereas Experiment 1.1 and 1.2 showed opposite patterns (see 10, Experiment 1.1 and 2 show very similar patterns, i.e. both show a difference in the fronting condition as well as a general trend in extraction.
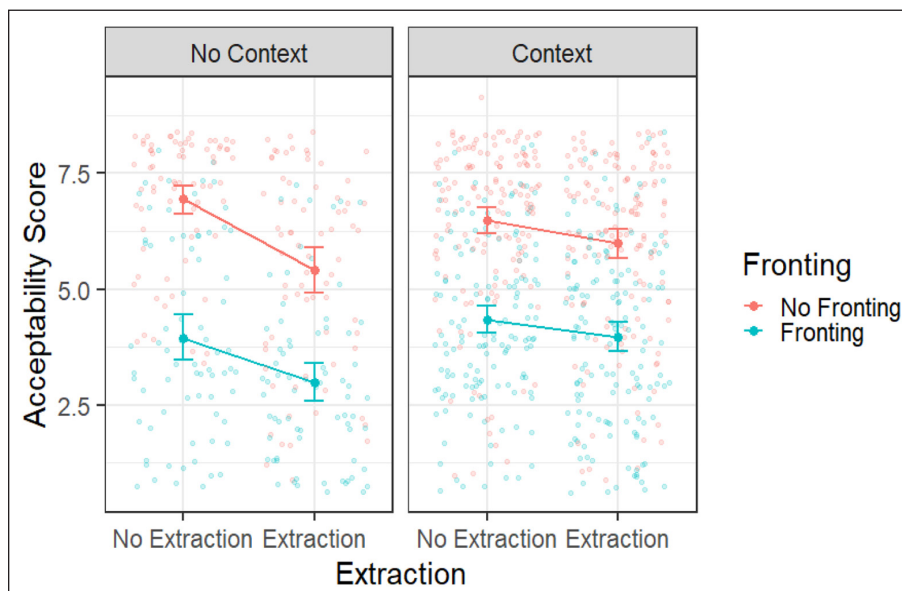
Figure 15 Results experiment 1.1 and 2. Error bars show the 95% confidence intervals.

Once again a mixed effects model was run similar to the one in Experiment 1.2: For the fixed effects, extraction, fronting, context and the interactions between the different variables were added to the model. As random effects, by-subject and by-item random intercepts of all main effects were added to the model. The model comparison shows that both fronting and extraction improved the model in comparison to a model in which these variables were excluded. The full model (AIC 3363.8) significantly improved the model compared to a model without fronting (AIC 3711.7) (χ^2 (1, N = 864) = 349.97, $p < 2.2e$–16): Fronted sentences are less acceptable than non-fronted ones (b = –2.844, t(806.58) = –13.16, $p < 2e$–16).

The full model (AIC 3363.8) was also an improvement to a model without extraction (AIC 3395.2) (χ^2 (1, N = 864) = 33.42, p = 7.421e–09): Sentences with wh-extraction are less acceptable than sentences without (b = –1.387, t(808.82) = –6.418, p = 2.34e–10).

More interestingly, there is a significant interaction effect between fronting and context. The full model (AIC 3345.3) was a significant improvement to a model without this interaction (AIC 3351) (χ^2 (1, N = 864) = 7.691, p = 5.55e–3).

Figure 16 in which Experiments 1.1 and 2 are plotted without *extraction*, clearly shows the interaction effect. The interaction effect tells us that adding a context significantly improves the *Fronting* condition, but not the *No-fronting* condition (b = 0.635, t(806.32) = 2.772, p = 5.71e–3). This result is compatible with a pragmatic explanation to freezing. (b = –0.495, t(56.918) = –1.168, $p > $ .05).

A similar significant interaction effect was found between *extraction* and *context*. The full model was a significant improvement to a model without this interaction (AIC 3358.7) (χ^2 (1, N = 864) = 15.4, p = 8.701e–05).
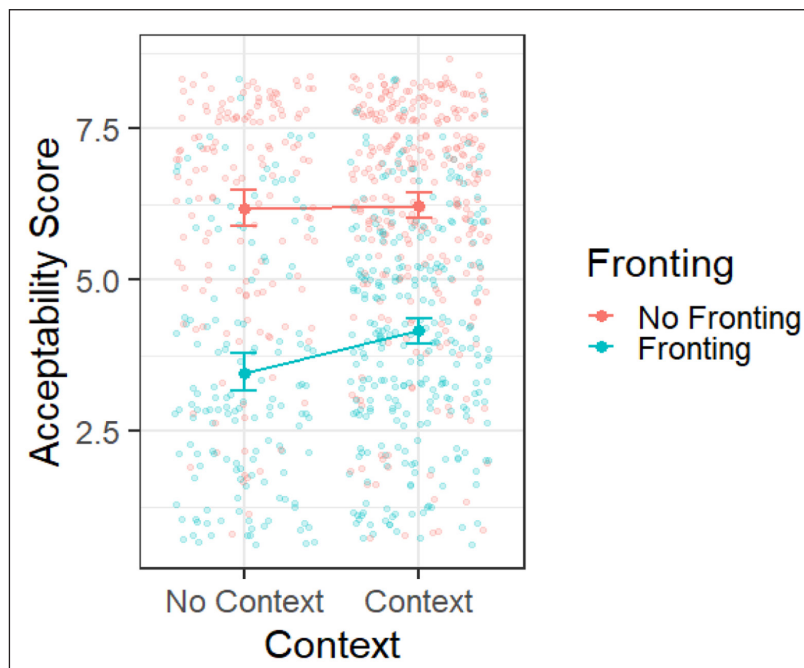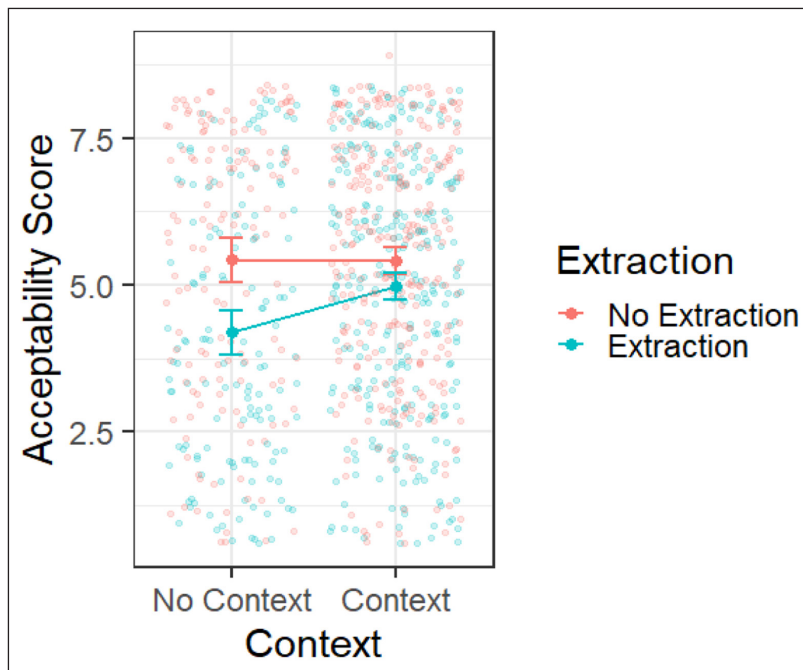

Figure 16 Results experiment 1.1 and 2 plotted without Extraction. Error bars show the 95% confidence intervals.

Figure 17 in which Experiments 1.1 and 2 are plotted without *fronting*, clearly shows the interaction effect. The interaction effect tells us that adding a context significantly improves the *Extraction* condition, but not the *No-Extraction* condition (b = 0.9186, t(818.22) = 3.940, p = 8.86e–05).

5 Conclusion and Discussion

Experiment 1.1 has shown us that the freezing effect can be described as an additive effect of *extraction* and *fronting*. This supports the intuitions of previous studies on different freezing constructions like PP-extraposition (Hofmeister et al. 2015) and Heavy NP shift constructions (Konietzko et al. 2018). These results indicate that freezing is not a phenomenon on its own, but the combination of several syntactic processes coming together.

Important to keep in mind is that evidence for additivity is based on finding main effects but no interaction effects between certain syntactic processes. An interaction effect indicates that a certain syntactic phenomenon cannot simply be explained by the effects of the individual syntactic processes involved, instead it is the combination of these processes that gives rise to a certain phenomenon. However not finding interaction effects does not mean that there really are not any interaction effects. For instance, it could be that the experiment did not have enough power to detect such an interaction effect. On the other hand, this seems unlikely since similar results have been found for three different freezing phenomena: PP extraction (Hofmeister et al. 2015), heave NP shift (Konietzko et al. 2018) and in the current study for DP-fronting. Additivity thus seems to be a robust effect.

Since both of these syntactic processes, i.e. wh-extraction and *movement*, have been linked to higher processing costs by previous processing studies, Hofmeister et al. (2015) and Konietzko et al. (2018) came to the conclusion that freezing is a matter of processing complexity rather than a structural constraint. Although additive effects were found in Experiment 1.1, the current study does not go as far as to assume that the freezing effect is therefore solely a matter of processing, nor does it exclude a processing explanation. The aim of the current study is not to provide evidence for a processing account, the aim is to provide evidence for pragmatic factors to freezing and argue against a prosodic account. What the comparison with Experiment 1.2 tells us is that at least part of the freezing effect, i.e. DP-fronting, can be explained in terms of pragmatic factors.

Culicover & Winkler (2010) and Winkler et al. (2016) argue for a prosodic account of freezing. The idea is that the freezing effect is the result of a prosodic garden path created by a conflict between the default prosody and the information structure encoded in the syntax. More specifically it is claimed that extraction can only take place from a focused constituent. As we have seen, Winkler et al. (2016)'s experimental design contained some limitations. First of all, even though Culicover & Winkler (2010) and Winkler et al. (2016) argue for a prosodic account, they did not look into production data at all. Rather the conclusions were based indirectly on acceptability judgments. To directly test the prosodic account, both experiment 1.1 and 1.2 collected participants' production data. The other limitation in the experimental design by Winkler et al. (2016) is that their stimuli were not designed to test the prosodic account specifically since the experiment could also be testing the presence or absence of a contextual anchor. Experiments 1.2 and 2 were designed to tease the prosodic account apart from the pragmatic account. The contrastive focus variable was created to specifically test the prosodic account and as we have seen, the results do not support a prosodic account. Although significant in Experiment 1.2, neither in Experiments 1.2 nor 2 did focus have the expected effect on the acceptability of the sentence based on a prosodic account of freezing. One would

have expected that the *DPFocus* context would significantly enhance fronting, whereas the *VPFocus* condition would not. There is however an overall effect of context, adding a context significantly improves both DP-fronting and wh-extraction.

This still did not provide conclusive evidence that the prosodic account is incorrect since the subjects could have produced completely different prosodic contours not reflecting the expected patterns. However, the production data from Experiment 1.2 showed us that the participants did produce the expected prosodic patterns. More specifically, it showed that even if the moved constituent was prosodically focused, it did not make the freezing cases more acceptable, therefore rendering the prosodic account less plausible. In addition, Experiment 2 showed that even when the task was simplified and the subjects perceived the right prosody, there was still no effect of focus prosody on their acceptability judgments.

The comparison between Experiment 1.1 and 2 provides evidence showing that pragmatics plays a role in explaining the freezing effect. By keeping the target stimuli across the different experiments the same and therefore keeping their syntax the same, one can specifically see the effects of adding a context to these sentences with fronted DPs. Comparing Experiment 1.1 with 2 showed us that adding a context significantly improves sentences with a fronted DP. One shortcoming of the current study however is that it is not exactly clear what part of the context improves DP-fronting nor is it clear how the context improves DP-fronting. The improvement could either come from the pre-sentence or the context since both contain a contextually given referents. In section 3.2.2 it was explained how in the pre-sentence a contrastive item is introduced which could function as a contextual anchor to the fronted DP. What was not taken into account was that all fronted DPs are also referentially given in the context (see Appendix C).

Explanations for why DP-fronting would improve given a contextually given referents, are plenty and diverse. If taking Rooth (1985)'s or Büring (1999)'s point of view according to which focus (or topic) invoke alternatives, then the improvement of DP-fronting could be explained in terms of the contextually given referents accommodating the set of alternatives. In other words, the contextually given referents is an element in the set of alternatives invoked by the focus/topic function of the fronted DP thus making it easier to associate the fronted DP with the contextually given referents.

On a different note, one could regard fronted DPs as D-linked elements. This makes sense since as we have seen in section 1.4, fronted DPs are dependent on previously mentioned elements in the discourse. D-linked elements require a higher level of initial activation in working memory compared to non D-linked elements, given their more referential nature (Goodall 2015). There is considerable evidence that this higher activation in working memory creates a processing advantage for D-linked elements when processing filler-gap dependencies and superiority effects (Kluender 1998; Frazier & Clifton 2002; Diaconescu & Goodluck 2004; Hofmeister 2011; Hofmeister et al. 2013). It is reasonable to assume that due to this higher activation in working memory, the association between a contextual anchor and fronted DP is more easily established which could result in higher acceptability.

On a similar note, the significant effect that context has on wh-extraction could be easily explained in terms of D-linked elements as well. All stimuli in this study consist of so called D(iscourse)-linked wh-words like *which* ..as opposed to non D-linked wh-words like *what* and *who*. It makes sense that these D-linked elements require a contextual referent, thus explaining the significant effect. It would be interesting to see if a similar significant effect can be found for non D-linked wh-words.

Notice how the pragmatic and processing account do not necessarily rule out one another, rather they seem to complement each other. It should become clear that a pragmatic phenomenon like the necessity of a contextual anchor can be explained from different perspectives that complement each other without having to choose one explanation over another. Important to note is that the aim of this study never was to propose a completely new analysis nor argue for one specific analysis, rather the aim was to provide empirical evidence showing that pragmatic factors play a significant role in explaining the freezing effect. Regardless of whether the improvement of fronted DPs comes from the pre-sentence or the context and regardless of how to account for this observation, the current research has shown that a contextually given referents significantly improves DP-fronting.

Although this study has only looked at embedded fronting constructions and their need for a contextual anchor, other constructions might have similar pragmatic requirements. The pragmatic function of extraposition for instance, has also been associated with definiteness, focus and topicalization (Bolinger 1992; Huck & Na 1990; Michaelis & Lambrecht 1996). Therefore it would not be surprising if similar pragmatic constraints played a role in the freezing effect found in extraposition cases.

Unfortunately, the pragmatic approach cannot account for the whole freezing phenomenon. The effect sizes that context has on for instance extraction and fronting are relatively small, so the freezing effect cannot be 100% explained in terms of a missing contextual anchor. Nonetheless, it can be concluded that the pragmatic account is part of the story explaining the freezing effect at least for frozen sentences involving DP-fronting. It thus seems that the freezing effect cannot simply be explained in terms of one account, rather different factors seem to play a role: these factors could be either constraints imposed by the grammar, processing difficulties or yet another factor. So instead of arguing for one account over the other, this study contributes to the debate by showing that at least pragmatic constraints play an important role in explaining the freezing effect.

Although this study acknowledges that pragmatic factors alone are not enough to explain the complete freezing effect and therefore other factors must be involved as well, the current research does show that the data cannot be explained by a prosodic account. In other words pragmatic factors contribute partially to the freezing effect, but other factors definitely play a role as well. However, when comparing the prosodic account against the pragmatic account the current study has made clear that the former account is not supported by the data and therefore the pragmatic account has more empirical coverage compared to a prosodic account. Which other factors contribute to the freezing effect and how much pragmatic factors play a role in the freezing effect compared to other factors is still left unanswered. This research has only been part of a small amount of experimental approaches to freezing and more research has to be done in order to get a more complete image of how these different factors interact with each other.

So far, it seems that a combination of structural factors, processing factors and pragmatic factors should be considered to account for the freezing effect. More research needs to be done to test how many factors play a role in accounting for the apparent freezing effect and how the different factors interact with each other.

Additional files

The additional files for this article can be found as follows:

- **Appendix A.** List of Standard Abbreviations. DOI: *https://doi.org/10.16995/glossa.5870.s1*

- **Appendix B.** Statistical Models. DOI: *https://doi.org/10.16995/glossa.5870.s2*

- **Appendix C.** Stimuli. DOI: *https://doi.org/10.16995/glossa.5870.s3*

Competing interests

The author has no competing interests to declare.

Author affiliation

Gouming Martens *orcid.org/0000-0002-1884-0759*
McGill University, CA

References

Bates, Douglas, Martin Maechler, Ben Bolker, Steven Walker, Rune Christensen, Henrik Singmann, Bin Dai, Fabian Scheipl, Gabor Grothendieck, Peter Green, John Fox, Alexander Bauer & Pavel Krivitsky. 2014. lme4: Linear mixed-effects models using eigen and s4. *R package version* 1(7). 1–23. DOI: *https://doi.org/10.18637/jss.v067.i01*

Bayer, Josef. 2005. Was beschränkt die extraktion?: Subjekt objekt vs. topic fokus. In Franz d'Avis (ed.), *Deutsche syntax: Empirie und theorie*, 233–257. Göteborg, Sweden: Acta Universitatis Gothoburgensis.

Bayer, Josef. 2012. From modal particle to interrogative marker: a study of german denn. In Laura Brugè (ed.), *Functional heads* 7. 13–28. New York, USA: Oxford University Press. DOI: *https://doi.org/10.1093/acprof:oso/9780199746736.003.0001*

Bayer, Josef & Hans-George Obenauer. 2011. Discourse particles, clause structure, and question types. *The linguistic review* 28(4). 449–491. DOI: *https://doi.org/10.1515/tlir.2011.013*

Birner, Betty & Gregory Ward. 1998. *Information status and noncanonical word order in english* 40. John Benjamins Publishing. DOI: *https://doi.org/10.1075/slcs.40*

Boeckx, Cedric & Kleanthes Grohmann. 2004. Submove: Towards a unified account of scrambling and d-linking. In *Peripheries*, 241–257. Springer. DOI: *https://doi.org/10.1007/1-4020-1910-6_10*

Boersma, Paul & David Weenink. 2002. Praat, a system for doing phonetics by computer. *Glot international* 5.

Bolinger, Dwight. 1992. The role of accent in extraposition and focus. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 16(2). 265–324. DOI: *https://doi.org/10.1075/sl.16.2.03bol*

Büring, Daniel. 1999. Topic. In Peter Bosch & Rob van der Sandt (eds.), *Focus — linguistic, cognitive, and computational perspectives*, 142–165. Cambridge University Press.

Chomsky, Noam. 1964. *Current issues in linguistic theory* 38. Mouton.

Chomsky, Noam. 2001. Derivation by phase (mitopl 18). *Ken Hale: A Life is Language*, 1–52.

Corver, Norbert. 2006. Freezing effects. *The Blackwell companion to syntax* 2. 383–406. DOI: *https://doi.org/10.1002/9780470996591.ch28*

Crain, Stephen & Janet Dean Fodor. 1985. How can grammars help parsers. David Dowty, Lauri Kartunnen & Arnold Zwicky (eds.), *Natural language parsing: Psycholinguistic, computational, and theoretical perspectives*, 94–128. DOI: *https://doi.org/10.1017/CBO9780511597855.004*

Culicover, Peter & Susanne Winkler. 2010. Freezing: A conspiracy. Ms., Ohio State University and Universität Tübingen.

Diaconescu, Rodica Constanta & Helen Goodluck. 2004. The pronoun attraction effect for d (iscourse)-linked phrases: Evidence from speakers of a null subject language. *Journal of Psycholinguistic Research* 33(4). 303–319. DOI: *https://doi.org/10.1023/B:JOPR.0000035103.19149.dc*

Emonds, Joseph E. 1976. *A transformational approach to english syntax: Root, structurepreserving, and local transformations*. Academic Press Inc.

Emonds, Joseph Embley. 1970. *Root and structure-preserving transformations*. Boston, MA, USA: Massachusetts Institute of Technology.

Frazier, Lyn. 1987. Sentence processing: A tutorial review. *Attention and Performance*.

Frazier, Lyn & Charles Clifton. 2002. Processing "d-linked" phrases. *Journal of Psycholinguistic Research* 31(6). 633–659. DOI: *https://doi.org/10.1023/A:1021269122049*

Goodall, Grant. 2015. The d-linking effect on extraction from islands and non-islands. *Frontiers in psychology* 5. 1493. DOI: *https://doi.org/10.3389/fpsyg.2014.01493*

Gregory, Michelle & Laura Michaelis. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of pragmatics* 33(11). 1665–1706. DOI: *https://doi.org/10.1016/S0378-2166(00)00063-1*

Gundel, Jeanette. 1977. Role of topic and comment in linguistic theory.

Gundel, Jeanette. 1988. Universals of topic-comment structure. *Studies in syntactic typology* 17(1). 209–239. DOI: *https://doi.org/10.1075/tsl.17.16gun*

Gundel, Jeanette & Thorstein Fretheim. 2004. Topic and focus. *The handbook of pragmatics* 175(196). 12. DOI: *https://doi.org/10.1002/9780470756959.ch8*

Hietaranta, Pertti. 1984. A functional note on topicalization. *English Studies* 65. DOI: *https://doi.org/10.1080/00138388408598302*

Hofmeister, Philip. 2011. Representational complexity and memory retrieval in language comprehension. *Language and cognitive processes* 26(3). 376–405. DOI: *https://doi.org/10.1080/01690965.2010.492642*

Hofmeister, Philip, Florian Jaeger, Inbal Arnon, Ivan Sag & Neal Snider. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28(1–2). 48–87. DOI: *https://doi.org/10.1080/01690965.2011.572401*

Hofmeister, Philip & Ivan Sag. 2010. Cognitive constraints and island effects. *Language* 86(2). 366. DOI: *https://doi.org/10.1353/lan.0.0223*

Hofmeister, Philip, Peter Culicover & Susanne Winkler. 2015. Effects of processing on the acceptability of "frozen" extraposed constituents. *Syntax* 18(4). 464–483. DOI: *https://doi.org/10.1111/synt.12036*

Huck, Geoffrey J. & Younghee Na. 1990. Extraposition and focus. *Language*, 51–77. DOI: *https://doi.org/10.1353/lan.1990.0023*

Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Edinburgh, UK: University of Edinburgh dissertation.

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. *Syntax and semantics*, 241–280. DOI: *https://doi.org/10.1163/9789004373167_010*

Konietzko, Andreas, Susanne Winkler & Peter Culicover. 2018. Heavy np shift does not cause freezing. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 1–11. DOI: *https://doi.org/10.1017/cnj.2017.56*

Koster, Jan. 1978. Why subject sentences don't exist. *Recent transformational studies in European languages*, 53–64.

Michaelis, Laura & Knud Lambrecht. 1996. Toward a construction-based theory of language function: The case of nominal extraposition. *Language*, 215–247. DOI: *https://doi.org/10.2307/416650*

Müller, Gereon. 2010. On deriving ced effects from the pic. *Linguistic Inquiry* 41(1). 35–82. DOI: *https://doi.org/10.1162/ling.2010.41.1.35*

Prince, Ellen. 1981. Topicalization, focus-movement, and yiddish-movement: A pragmatic differentiation. In *Annual meeting of the berkeley linguistics society* 7. 249–264. DOI: *https://doi.org/10.3765/bls.v7i0.2092*

Prince, Ellen F. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In *The limits of syntax*, 281–302. Brill. DOI: *https://doi.org/10.1163/9789004373167_011*

R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. *https://www.R-project.org/*.

Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica anc Studia Philosophica Gandensia* 27(1). 53–94.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, 281–337. Springer. DOI: *https://doi.org/10.1007/978-94-011-5420-8_7*

Rooth, Mats. 1985. *Association with focus*. Amherst, MA, USA: University of Massachusetts, Amherst dissertation.

Ross, John. 1967. *Constraints on variables in syntax*. Boston, MA, USA: Massachusetts Institute of Technology dissertation.

Ross, John. 1974. Three batons for cognitive psychology. *Cognition and the symbolic processes*.

Schütze, Carson. 2016. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Language Science Press. DOI: *https://doi.org/10.26530/OAPEN_603356*

Sprouse, Jon. 2007. Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics* 1. 123–134.

Vallduví, Enric. 1993. The informational component. *IRCS Technical Reports Series*, 188.

Van Riemsdijk, Henk C. 1978. A case study in syntactic markedness: The binding nature of prepositional phrases. *Studies in generative grammar* 4. 1–313.

Ward, Gregory & Betty Birner. 1994. A unified account of english fronting constructions. *University of Pennsylvania Working Papers in Linguistics* 1(1). 11.

Ward, Gregory & Betty Birner. 2004. Information structure and non-canonical syntax. *The handbook of pragmatics*, 153–174.

Ward, Gregory, Betty Birner & Rodney Huddleston. 2002. *The cambridge grammar of the english language* chap. Information packaging, 1363–1447. Cambridge University Press. DOI: *https://doi.org/10.1017/9781316423530.017*

Wexler, Kenneth & Peter Culicover. 1980. Formal principles of language acquisition. *Language*.

Winkler, Susanne, Janina Radó & Marian Gutscher. 2016. What determines 'freezing'effects in was-für split constructions? *Quantitative Approaches to Grammar and Grammatical Change: Perspectives from Germanic* 290. 207.