A heuristic mathematical approach for modeling constraint cumulativity: Contrastive focus in Spanish and Catalan

ARIA ADLI

Abstract

This work presents a heuristic mathematical approach to modeling and quantifying constraint violations. The empirical phenomenon consists of different word orders with contrastive focus in the Spanish variety of Catalonia and in Catalan. Experimental data has been collected with a gradient acceptability judgment test for auditorily presented focus constructions.

The most unmarked position for contrastively focused objects in the Spanish of Catalonia is in situ or sentence-final, while in Catalan sentence-initial and in situ position is preferred. In Spanish, the subject has to be postverbal if the focus is preposed, but these constructions still remain marked.

These results are explained with a model according to which each constraint has a numerical violation cost, and according to which the costs of all constraints violated by a sentence cumulate. The presumed constraint set is analyzed with an algebraic approach based on a system of linear equations with which violation costs are calculated. The methodology is heuristic in the sense that it can be used for exploring whether there could be a "missing piece". Such a missing piece would be another constraint to be included post hoc in the set. It corresponds to another unknown added to the system of linear equations.

1. Introduction

In this work I want to present a heuristic mathematical approach and discuss its possibilities as well as its limitations in modeling and quantifying constraint violations. Gradient acceptability judgments have been collected on various constructions with contrastive focus in Catalan and the Spanish variety of Catalonia, two typologically and geographically closely related languages. Subjects

The Linguistic Review 28 (2011), 111–173
DOI 10.1515/tlir.2011.004
0167–6318/11/028-0111

listened to recordings of test sentences, which systematically varied with respect to two structural aspects: the position of the focused element, and preverbal versus postverbal subject (the test sentences will be presented in Section 2).

The data is analyzed in a model of cumulative constraint violation costs. Different models of constraint cumulativity have been proposed in the literature. The one I apply is probably closest to Keller's (2000) approach (which has inherited important elements from Uszkoreit 1987; Legendre et al. 1990; and has probably also influenced Potts et al. 2010). The heuristic mathematical approach can provide new theoretical insights, because it can hint at the relevance of constraints that had not been integrated in the initial tableau. At the same time, it is not an automatic problem solver: In some cases it can be applied seamlessly to the data. In others its application demands compromises, e.g., a reformulation of one of the constraints of the initial set, and the researcher is then required to decide whether she/he accepts such modification.

The primary goal of the heuristic mathematical methodology which is constitutive of the subfield of experimental mathematics, is not hypothesis testing but hypothesis generation, i.e. it is used to gain conceptual insights in model building. Heuristic quantitative techniques are commonly used in psychology and in the social sciences, manifest for example in studies that apply factor analysis, cluster analysis or canonical correlation. This approach is not new in linguistics either, although it has not been linked, as far as I know, to this methodological context. Quantitative simulation studies fall into this class, for example the simulations conducted by Jäger (2007) in his evolutionary game theory model to typological variation in case systems.

I am not proposing a new theory of constraint competition. Rather, I present a complementary methodological approach, which can be useful in modeling acceptability judgments. It can prove useful in studies that try to relate theoretical and empirical perspectives in linguistics – a line of research that has been enjoying increased interest in recent years (see for example the papers in Featherston and Winkler 2009; Winkler and Featherston 2009; Boye and Engberg-Pedersen 2010). I will leave the question how this approach should be modified in order to be applicable to frequency data from corpora, for future research. Furthermore, I am not concerned with the issue of probabilistic modeling, such as stochastic OT (Boersma and Hayes 2001) or maximal entropy (Berger et al. 1996).

After laying out basic notions of contrastive focus constructions, especially for Romance, I present the experimental acceptability judgment test. Then I analyze the data in a series of analyses of variance (ANOVA) and formulate, based on these results, the initial set of constraints. Once this much background is established, the study turns to the heuristic mathematical approach. It is applied first on data from Catalan, then on data from the Spanish of Catalonia. The entire procedure will result in a set of constraints, each of them associated with a specific violation cost, which can explain the fine-grained differences in markedness between different word orders of focus constructions in both languages. The paper concludes with a discussion on the possibilities and limitations of the heuristic mathematical approach.

1.1. Nomenclature of focus

At the most general level, focus represents the non-presuppositional part of the sentence as opposed to the presuppositional part (Chomsky 1971; Jackendoff 1972). Zubizarreta and Vergnaud (2005: Sec. 2) suggest the definition in (1), an idea that goes back to Halliday (1967).

(1) In a sentence S = α X β, X is a possible focus just in case S would count as a felicitous answer to the question Q = wh-X $\alpha\beta$.

Krifka (2007) proposes a definition based on the insight provided by Alternative Semantics (Rooth 1985, 1992): "Focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions". He situates focus in the broader context of information structure and information packaging, departing from Chafe's (1976) proposal, according to which information packaging can be described as the continuous change of the Common Ground (Karttunen 1974) of the discourse participants. Various types of focus have been distinguished in the literature (under differing labels, though, which can sometimes be confusing).

First, we distinguish between different domains of the focus structure.[1] The domain of the focus in the answer is explicit in the corresponding question. *Wide focus* refers to sentences with the largest focus structure (the entire TP) which answer the question "what happened?" as is shown in (2). We use the term *narrow focus* to refer to all focus structures that correspond to a constituent below TP, e.g. the subject DP, the entire VP, or any of the VP arguments. (3) demonstrates a case in which the object argument is focused.

(2) Q. *What happened?*
A. [*John bought a car in Madrid*]$_F$.

(3) Q. *What did John buy in Madrid?*
A. *John bought [a car]$_F$ in Madrid.*

1. Unlike Zubizarreta (1998) and Zubizarreta and Vergnaud (2005), I prefer the term *domain* of focus over *scope* of focus, in order to refer to wide and narrow focus. It is important to be aware of the fact that the notion of scope does not coincide with the notion of focus (see constructions with focus-sensitive operators such as 'only', Rooth 1985).

Second, we distinguish between *i-focus* and *c-focus,* a distinction based on the semantic property of exhaustivity. Exhaustivity means that the description or referent has to be unique (Halliday 1967; Rochemont 1986; Kiss 1998). A focus constituent with this property "exhaustively identifies the proper subset of a contextually relevant set of entities as the one for which the predicate holds" (Horvath 2000: 201). Following Zubizarreta and Vergnaud's (2005) terminology, exhaustive focus is referred to as contrastive focus (or c-focus or c-F) and non-exhaustive focus as informational focus (or i-focus or i-F). Many languages show different surface patterns for c-focus compared to i-focus (see Selkirk 2002). For example, c-focus in Hungarian must appear left-adjacent to the verb, while i-focus can appear in various positions, typically postverbally (Kiss 1998). The two types have different contexts: a question in the case of i-focus and an assertion in the case of c-focus. For example (4a) is the context of the i-focus in (4b), which is in turn the context of the c-focus in (4c).[2] The distinction between i-focus and c-focus basically corresponds to what Kiss (1998) calls "presentational focus" and "identificational focus" (apart from the fact that according to him, the features [+exhaustive] and [+contrastive] do not co-occur in all languages).

(4) a. *What did John buy in Madrid?*
b. *John bought* [*a* <u>*car*</u>]$_{\text{i-F}}$ *in Madrid.*
c. *No, John bought* [*a BICYCLE*]$_{\text{c-F}}$ *in Madrid.*

This example also highlights another semantic difference between the two types of focus. While focus in general has the function of assigning a value to the variable introduced by the presupposition, c-focus, in addition, negates the value given to the variable by the presupposition in order to assign an alternative value to it (Zubizarreta 1999: 4226-4228). Semantically, both emphasis as in (6a) and c-focus as in (6b) make a statement about the truth-value of the assertion introduced by its context statement in (5). However, c-focus also assigns a different value to the variable introduced in (5). Consequently, emphasis has a purely metagrammatical function (correction, repair, denial, reassertion of the hearer's presupposition), while c-focus has both a metagrammatical and a grammatical function (i.e., also the introduction of a variable with its value in establishing a focus-background structure). Further work on the semantic

2. Henceforth, I illustrate nuclear stress by underlining the entire word on which the nuclear accent (a term I reserve to i-focus) falls. Contrastive and emphatic stress is always represented by UPPERcase letters. The domain is marked by the index F (more precisely i-F or c-F) following the closing bracket, e.g., [...]$_{\text{i-F}}$. The symbol for the focus domain should not be confounded with the focus feature F which will be shown as an index preceding the closing bracket, e.g., [... $_{\text{F}}$].

aspect of focus has been carried out by Szabolcsi (1981), Krifka (1991, 2007), Diesing (1992), Rooth (1992), and Schwarzschild (1999).

(5) *I'm sure that John bought a car in Madrid, although he has no money.*

(6) a. *You are right. John DID buy a car in Madrid.*
b. *No, John did not buy a car. He bought a BICYCLE in Madrid.*

Third, we distinguish between *expression focus* and *denotation focus* (Krifka 2007). The former is typically used for corrections, often correlating with negation, and it does not need to carry constituents or meaningful units, as illustrated by (7).

(7) *No, you don't write ex*$[TRO]_F$*verted but ex*$[TRA]_F$*verted.*

However, the great majority of the literature is concerned with denomination focus, which carries on the meaning. Technically, it concerns the meaning $\|\alpha\|$ of an expression, where focus leads to the assumption of a set of alternative meanings. For example, sentence (4b) ('John bought a $car_{i\text{-}F}$ in Madrid') induces the alternatives $\{bought(car)(x)|x \in entity\}$. Based on this distinction, Krifka (2007) proposes the definition given in (8) with which I work. It is inspired by Rooth's (1992) Alternative Semantics. The focus-induced alternatives correspond to the Hamblin (1958, 1973) meaning of questions, i.e., a set of propositions, each being the denotation of a congruent answer (see Krifka 2004 for further discussion).

(8) A property F of an expression α is a Focus property iff F signals (a) that alternatives of (parts of) the expression α or (b) alternatives of the denotation of (parts of) α are relevant for the interpretation of α.

Not all combinations of these three dimensions given above are possible. C-focus cannot be wide but only narrow. Furthermore, expression focus cannot be i-focus, but only c-focus (and thus only narrow). Building on this nomenclature, we can narrow down the phenomenon we are studying in the experimental sentences to *narrow denotation c-focus in declaratives*.[3]

1.2. Nuclear stress and prosodically motivated movement in Romance

In the following I present some preliminaries on the relation between prosody, focus, and syntax. For reasons of clarity, I will first present some facts con-

3. Some languages exhibit syntactic configurations that correspond to *contrastively* focused *wh*-questions. This is for example the case with scrambled *wh*-NPs in Persian (Karimi 2005) or with *wh*-clefts in French (Zubizarreta and Vergnaud 2005). Unlike in declaratives, c-focus questions do typically not negate the previous value given to the variable by the presupposition. Rather, they have a *strong* existential presupposition (Adli 2010).

cerning i-focus before turning to c-focus. Certain positions of nuclear stress are incompatible with certain focus domains. Observe the contrast between (9a) and (9b). It can be explained by the principle in (10) which is the *Nuclear Stress Rule* (abbreviated NSR), more precisely its version proposed for Romance languages (Zubizarreta 1998; Zubizarreta and Vergnaud 2005).[4] Its definition incorporates the intonational phrase or I-phrase of the prosodic hierarchy (Liberman 1975; Pierrehumbert 1980).

(9) a. *[*El* <u>*bebé*</u> *llora*]$_{\text{i-F}}$. (Sp.)
the baby cries
b. [*El bebé* <u>*llora*</u>]$_{\text{i-F}}$. (Sp.)

(10) In Romance, nuclear stress falls on the rightmost accented word within the I-phrase.

When prosodic means are used for information packaging, in particular for distinguishing between focus and background, then there must be some sort of constraint on the placement of prosodic prominence, which is the NSR for Romance in (10), and some sort of constraint on the correspondence of focus and prosody, which is given in (11). The latter goes back to Chomsky (1971) and Jackendoff (1972), and is known since Zubizarreta (1998) under the label *focus-prosody correspondence principle* (FPCP). Furthermore, Zubizarreta (1998: 21) proposes the *Focus Prominence Rule* (FPR) in (12), which builds on the FPCP. The FPR captures, in X-bar theoretic terms, the fact that the F-structure of the sentence is constrained by the location of main prominence (and not by the distribution of the pitch accents, as claimed by Gussenhoven 1984; Selkirk 1984, 1995).[5]

4. The Romance version in (10) is very similar to the first formulation of the NSR proposed by Chomsky & Halle (1968) for Germanic languages ('The rightmost word-level stress of a phrase bears main stress within that phrase'). Unlike Romance, subsequent work showed that the algorithm for NSR is more complicated in Germanic, because the placement of nuclear stress is sensitive to syntactic structure in these languages. In order to capture this sensitivity, Zubizarreta (1998: 56) proposed a modularized NSR in two parts where the first one applies to a tree structure without the deaccented nodes, and the second one to the full-fledged tree based on asymmetric c-command. However, Romance languages, except for French (and also except for Catalan, see below), do not deaccent anaphoric and defocalized material, which means that there is no need to distinguish between syntactic and metrical tree (Zubizarreta 1998: 77).
5. The F-structure is a syntactic structure in which the constituent interpreted as focus or part of the focus has been marked with the diacritic [F] (following Jackendoff 1972, the F-structure contains additional information on the division between focus, i.e., [+F], and presupposition, i.e., [−F]).

(11) FPCP: The focused constituent (or F-marked constituent) must contain the intonational nucleus of the intonational phrase, where the intonational nucleus is identified as the syllable that bears main phrasal prominence.

(12) FPR: Given two sister categories C_i (marked [+F]) and C_j, (marked [−F]), C_i is more prominent than C_j.

Both principles together, the FPCP (11) and the NSR for Romance (10), predict that the focused element must occur at the right edge of the corresponding I-phrase. Zubizarreta (1998) highlights the resulting syntactic consequences. Any defocalized constituent that appears to the right of the i-focused element within the same I-phrase has to be displaced, as shown in the illicit case (14a) (the configuration in (14a) would only be licit with c-focus). This can be achieved either by placing the defocalized element after an intonational boundary, i.e., in a different I-phrase, as is shown in the right-dislocation in (14b). Or it can be done by *p*-movement as in (14c). Like scrambling (and also Scandinavian object-shift), *p*-movement targets an intermediate clause position (see López 2009: 8–11 on the difference between scrambling and p-movement).

(13) *Qué comió María en el restaurante?* (Sp.)
what ate Maria in the restaurant

(14) a. **María comió* [*una* $\underline{\textit{sopa}}$]$_{\text{i-F}}$ *en el restaurante.* (Sp.)
Maria ate a soup in the restaurant

b. *María comió* [*una* $\underline{\textit{sopa}}$]$_{\text{i-F}}$ # *en el restaurante.* (Sp.)
Maria ate a soup in the restaurant

c. *María comió en el restaurante* [*una* $\underline{\textit{sopa}}$]$_{\text{i-F}}$. (Sp.)
Maria ate in the restaurant a soup

By displacing the defocalized constituent, nuclear stress can fall on the rightmost constituent within the I-phrase. In these examples, the object-DP carries narrow i-focus. This type of movement is prosodically motivated in order to comply with both the FPR and the NSR. Therefore Zubizarreta calls the displacement of the entire [−F]-marked material into a position above the [+F]-marked constituent *p*-movement. The purpose of this movement operation is not feature checking in the Minimalist sense. Its purpose is "that the focalized constituent is in a position to receive prominence via the [NSR] thus ensuring that the output is compatible with the FPR" (Zubizarreta 1998: 124).

The same is claimed for prosodically prominent subjects. According to Zubizarreta (1998: 76, 127), an i-focus construction like (15a), whether [$\underline{\text{S}}$]$_{\text{i-F}}$VO or [$\underline{\text{S}}$VO]$_{\text{i-F}}$, is not possible, because it violates the NSR in (10). According to Zubizarreta (1998: 127) the problem is resolved by moving all [−F]-marked material to a higher position, in order for the prosodically prominent subject to

AUTHOR'S COPY | AUTORENEXEMPLAR

appear in a postverbal position. With respect to (15b) with its bracketed representation in (15c), she first assumes overt verb movement to T driven by a strong V-feature in T, followed by the adjunction of the entire VP to the left of *v*P. She proposes that this second step is *p*-movement. Note that I do not follow the assumption defended by some authors that VOS is the underlying word order in Catalan (Bonet 1990; Vallduví 1993) and Spanish (Fernández-Soriano 1989; Contreras 1991).

(15) a. *El <u>niño</u> come el helado.* (Sp.)
the boy eats the ice-cream

b. [*Come el helado el <u>niño</u>*]$_{\text{i-F}}$ (Sp.)
eats the ice-cream the boy

c. [$_{\text{TP}}$ *come T* [$_{v\text{P}}$ [$_{\text{VP}}$ *un helado*]$_{\text{i}}$ [$_{v\text{P}}$ *el niño*$_{[+\text{F}]}$ *v* ~~*come*~~ *t*$_{\text{i}}$]]] (Sp.)

However, Gabriel (2007: 172), based on recorded speech data, shows that a default postverbal subject position is restricted to a specific case, namely to i-focus constructions with a cliticized object, e.g. *lo* substituting *el helado* in (16b). As soon as the object is a full DP, the default subject position is preverbal, thus (15a) > (15b) ($P > Q$ means that P is preferred over Q). This preference is supposed to be even stronger in ditransitive constructions, i.e., (17a) > (17b). A similar situation has been reported for Catalan (López 2009: 132), where the basic word order of wide i-focus is SVO as in (18).

(16) a. [*Come el helado el niño*]$_{\text{i-F}}$ (Sp.)

b. [*Lo come el niño*]$_{\text{i-F}}$ (Sp.)
CLIT$_{\text{3sg}}$ eats the boy

(17) a. [*El niño da el helado a su hermana*]$_{\text{i-F}}$ (Sp.)
the boy gives the ice-cream to his sister

b. [*Da el helado a su hermana el niño*]$_{\text{i-F}}$ (Sp.)
gives the ice-cream to his sister the boy

(18) [(*que*) *el Joan finalment va portar els llibres*]$_{\text{i-F}}$ (Cat.)
(that) the Joan finally brought the books

1.3. C-focus in Romance

The above-mentioned debate on the nature of preverbal subjects relates to the issue whether in general fronted focused elements in sentence-initial position (*foco antepuesto*) are necessarily contrastive in Spanish as in the [S]$_{\text{c-F}}$VO construction (19).

(19) [*El NIÑO*]$_{\text{c-F}}$ *come un helado.* (Sp.)
the child eats an ice-cream

The controversy in the literature is obvious. While some authors only consider c-focus possible (di Tullio 1997: 363; Zubizarreta 1998: 20–22, 1999: 4239; Ordóñez 2000: 29ff.; Revert-Sanz 2001: 27; Samek-Lodovici 2001: 347; Gutiérrez-Bravo 2002: 51; Zagona 2002: 211; Martín-Butragueño 2005: 135), some also observe fronted i-focus and would accept for example, a $[S]_{i\text{-}F}$VO variant of (19) (Domínguez 2004: 150).

Concerning the relation between prosody and syntax, example (19) has shown that contrastive focus cannot be subject to the NSR. One has to distinguish between two types of phrasal dominance related to contrastive and non-contrastive focus, which are presumably governed by different rules, a fact already pointed out by Chomsky (1971) and Jackendoff (1972). The purpose of this rule is to identify c-focus and metalinguistic functions such as emphasis (see Casielles-Suarez 1997 on Spanish; Prieto 2002a on Catalan). As regards c-focus, phrasal prominence is not generated by NSR, but by the Focus/Contrastive Stress Correspondence Principle in (20). In other words, there are two different main prominence assignment rules: the NSR and the Focus/Contrastive Stress Correspondence Principle (Zubizarreta 1998: 45, 77). The Catalan examples (21a) and (21b) show that the focus domain can be restricted to the contrastively stressed noun head. It cannot cover the whole subject DP because the included PP cannot be F-marked according to (20): N^0 does not dominate the PP (technically speaking, the focus projection rules postulated by Selkirk 1995: 561 in order to explain the domain of i-focus, do not apply; c-focus has much more restricted projection rules).

(20) A word with contrastive stress must be dominated by every F-marked constituent in the phrase.

(21) a. *El* $[$*CAFÈ*$]_{c\text{-}F}$ *de l'Ecuador és molt fort (i no el cacau de l'Ecuador).* (Cat.)
'The coffee from Ecuador is very strong (not the cocoa from Ecuador)'.

b. **El* $[$*CAFÈ de l'Ecuador*$]_{c\text{-}F}$ *és molt fort (i no el cacau del Brasil).* (Cat.)
'The coffee from Ecuador is very strong (not the cocoa from Brazil)'.

The placement of contrastive stress is less restricted than the placement of nuclear stress in i-focus. It can be realized on syllables that are metrically weak in the phonological word. Recall the examples of expression focus (7) above. Also recall the Spanish example (14a) above, showing that $O_{i\text{-}F}$ in an SVOP is infelicitous. The same order with $O_{c\text{-}F}$ is felicitous as (22) shows. Unlike i-focus, c-focus in Romance can also be realized in situ or in a fronted position.

(22) *María comió* [*UNA SOPA*]$_{\text{c-F}}$ *en el restaurante.* (Sp.)
Maria ate a soup in the restaurant

Vallduví (1995) points out that an important element in assuring focus prominence in Catalan is defocalization of the material following the c-focus. Defocalized material can be long, i.e. cover various phrases, as shown in the Catalan sentences (23), (24a), and (24b). Defocalized XPs are free to appear in different linear orders, which is not the case in constructions without c-focus and defocalization as in (25a) and (25b). Interestingly, examples (26a) and (26b) from Zubizarreta (1998: 156) show that the order of defocalized material after c-focus in Spanish is not free. She argues against Chomsky's (1981) and Rizzi's (1982) claim according to which postverbal subjects in Spanish (and other Romance languages) are right-adjoined to VP due to the fixed order of deaccented object-PP. Vallduví (1995) proposes that defocalized XPs in Catalan are not in their in situ position but left- or right-dislocated. According to his analysis, the preposed c-focus in (23), (24a), and (24b) is in situ, while all other constituents have been right-dislocated and adjoined to IP. While Vallduví (1993) applies his dislocation analysis to all preverbal subjects (for counter-arguments, see Sheehan 2007; Feldhausen 2008), I do not adopt it in the following.

(23) [*Al FUSTER*]$_{\text{c-F}}$ *la mare va donar les claus.* (Cat.)
to the carpenter the mother gave the keys

(24) a. [*A DISNEYWORLD*]$_{\text{c-F}}$ *portarem el nen aquest estiu.* (Cat.)
to Disneyworld take$_{\text{1pl,fut}}$ the boy this summer

b. [*A DISNEYWORLD*]$_{\text{c-F}}$ *el nen portarem aquest estiu.* (Cat.)
to Disneyworld the boy take$_{\text{1pl,fut}}$ this summer

(25) a. **Portarem a Disneyworld el nen aquest estiu.* (Cat.)
take$_{\text{1pl,fut}}$ to Disneyworld the boy this summer

b. **El nen portarem a Disneyworld aquest estiu.* (Cat.)
the boy take$_{\text{1pl,fut}}$ to Disneyworld this summer

(26) a. *Escondio* [*el NIÑO*]$_{\text{c-F}}$ *el libro debajo de la cama.* (Sp.)
hid the boy the book under the bed

b. **Escondio* [*el NIÑO*]$_{\text{c-F}}$ *debajo de la cama el libro.* (Sp.)
hid the boy under the bed the book

AUTHOR'S COPY | AUTORENEXEMPLAR

Note that certain elements are generally *deaccented* (Ladd 1980: 81; Gussenhoven 1984; Selkirk 1984), i.e., they are pronounced with a reduced pitch range, perceived as less prominent, and therefore cannot be the locus of nuclear stress (e.g., anaphoric elements in Germanic languages). An interesting difference between Catalan and Spanish is that defocalized material after c-focus in Catalan is always deaccented, pronounced with a reduced pitch on a low plateau. However, defocalized material after c-focus in Spanish does not necessarily need to be deaccented (Zubizarreta 1998: 156); see Selkirk (1995) on deaccented postfocal material.

The sentences above also show another interesting point: Preverbal c-focus in Catalan does not need to be left-adjacent to the verb, as we can see in (23) and (24b). Catalan behaves in this respect like Italian. Example (27) is taken from Rizzi (1997: 299) (the observation goes back to Antinucci and Cinque 1977). Such data is supporting the assumption of a rich CP-layer in Catalan. Rizzi's (1997) assumption of the full-fledged C-system is shown in (28). It includes an optional Focus or Emphasis projection, as well as an optional (recursive) Topic projection located between CP and TP.

(27)	[*QUESTO*]$_{\text{c-F}}$	*Gianni*	*ti*	*dirà*	*(non*	*quello*		(Ital.)
	this	Gianni	to-you	say$_{3,sg,fut}$	(not	that		
	che	*pensavi)*						
	what	thought$_{2,sg,past}$)						

(28) . . . C^0 (Top*) (Foc) (Top*) . . .

Barbosa (2001: 33) assumes that the CP layer in Spanish lacks FocP as an independent head and that *foco antepuesto* is therefore moved to the Spec of TP. Zubizarreta (1998: 100) presents a similar view in her *generalized TP analysis*, assuming that in a number of languages discourse-based functional features (topic, focus, emphasis, etc.) can combine with the T-feature to form syncretic categories such as T/topic, T/focus, etc. It has been argued that preverbal focus in Spanish is, unlike for example in Italian, supposed to be obligatorily left-adjacent to the verb (which is questionable, as will be discussed later). A structure where T is the landing site for several discourse-related elements, with topic in TopP preceding focus in T, would look as in (29).

(29) $[_{\text{CP}}$C $[_{\text{TopP}}$ Top* $[_{\text{TP}}$ T$_{\text{Foc/Wh}}$ $[_{\text{vP}}$ Subject v $[_{\text{VP}}$V Object]]]]]

However, we will see in Section 3.3 that the experimental results for the Spanish variety of Catalonia support a rich CP-layer. My working hypothesis is that both Catalan and the Spanish variety of Catalonia have a C-system as in (28).

1.4. Constraint-based approach

My point of departure are OT syntax approaches that have been proposed for focus-based word order variation, on which I will build a model based on cumulative constraint violation costs. A recent OT syntax work on focus and word order has been presented by Gabriel (2007: 236–255). He splits Grimshaw's (1997) OT-syntax translation of the basic rule of derivational economy (30) into (31a), and (31b) in order to reflect Zubizarreta's (1998) distinction between genuine syntactic movement and *p*-movement.

(30) STAY:
Do not move.

(31) a. *COPY:
No copies.
b. STAY-ϕ:
No chain-external material.

Zubizarreta (1998: 105–107) considers fronting of a focused element in Spanish as an operation of core syntax, on a par with *wh*-movement. Building on previous work on the specifics of focus movement (see Szendröi 2005 for an overview), I do not reduce movement of [+F]-marked and [−F]-marked material to the same constraint, thus splitting the "classic" STAY constraint into three parts: STAY-ϕ that applies to *p*-movement, *COPY$_{[-F]}$ that applies to the displacement of (non-*p*-moved) defocalized (or focally neutral) material, and *COPY$_{[+F]}$ that applies to the displacement of focused material. I leave the question open as to whether *COPY$_{[+F]}$ also applies to *wh*-movement.

(32) a. *COPY$_{[-F]}$:
No copies of [−F]-marked material.
b. *COPY$_{[+F]}$:
No copies of [+F]-marked material.

Furthermore, we will need Grimshaw and Samek-Lodovici's (1998: 194) formulation (33) of the *Full Interpretation* (FULLINT) constraint (originally introduced in Grimshaw 1997), and Gussenhoven's (2004: 160) ALIGNFOC constraint in (34) (originally introduced in Grimshaw and Samek-Lodovici 1998). I also refer to Büring and Gutiérrez-Bravo's (2001) FOCUSPROMINENCE constraint in (35), which reformulates Zubizarreta's (1998: 21) FPR in (12).

(33) FULLINT:
Parse lexical conceptual structure. Failed by expletives and auxiliary do.

(34) ALIGNFOC:
The focus constituent right-aligns with the I-phrase.

(35) FOCUSPROMINENCE:
Focus is most prominent.

Gussenhoven's (2004: 160) ALIGNFOC constraint (34), which goes back to Grimshaw and Samek-Lodovici's (1998) essentially identical ALIGNFOCUS constraint, expresses the reduced C-NSR for Romance in (10). Unlike C-NSR, which only applies to i-focus, (34) is supposed to apply to both i-focus and c-focus. Note that (34) belongs to the family of alignment constraints that concern the relation between grammatical and prosodic structure (see Truckenbrodt 2007: 437ff.).

Working with a stochastic OT approach, Gabriel (2007: 300) stresses that (31b), (33), and (34) are the three constraints on which an analysis of focus-based word order variation in Romance has to concentrate. Apart from the fact that they can cross-linguistically appear in a different default order, they can also appear in a different order within a language due to the probabilistic rules of the framework. However, their default ranking in Spanish is, according to him, as in (36). FOCUSPROMINENCE is not relevant for modeling variation because it is supposed to be the highest-ranked constraint, i.e., to be non-violable. He considers ALIGNFOC, FULLINT, and STAY-ϕ to have a common area of overlap in the sense of stochastic OT in order to explain different outputs, i.e., basically every possible ranking can actually occur (still, the ranking in (36) has the highest probability).

(36) ALIGNFOC $\gg$ FULLINT $\gg$ STAY-ϕ

(37) a. [*El* *niño* *come* *el* *helado*]$_{\text{i-F}}$ (Sp.)
the boy eats the ice-cream
b. [*Come el helado el* *niño*]$_{\text{i-F}}$ (Sp.)

As regards the examples (15a) and (16a), repeated as (37a) and (37b), FULLINT is violated in (37b) due to the postverbal subject. The postverbal subject is supposed to remain in *v*P-internal position (Suñer 1994) such that the EPP-feature in the Spec of TP has to be deleted by *pro*$_{\text{Expl}}$.[6] In addition, this construction violates *COPY due to the movement of the verb, and it violates STAY-ϕ due to *p*-movement of the direct object which becomes chain-external. Therefore the well-formedness of (37b) indicates that there must be at least one other higher-ranked constraint that is not violated by the sentence – which is precisely ALIGNFOC. On the other hand, an i-focus construction with preverbal subject such as (37a) violates neither FULLINT nor STAY-ϕ. However, it does violate ALIGNFOC because the subject-DP that carries nuclear stress is not

6. The underlying EPP definition is as follows: "The specifier of the highest functional head of the T-domain is filled by an overt element, by *pro* or by *pro*$_{\text{Expl}}$". Similar definitions have been proposed by Gutiérrez-Bravo (2006: 44, 73) and Fernández-Soriano (2004).

at the right edge of the I-phrase. (37a) also violates *COPY because the preverbal subject is raised to the Spec of TP. Assuming the ranking ALIGNFOC $\gg$ FULLINT $\gg$ STAY-ϕ, a standard OT approach would predict that (37a) is ungrammatical. However, both (37a) and (37b) are possible constructions. The stochastic OT approach can explain why the selection points can differ between two evaluations such that ALIGNFOC is sometimes lower ranked and other times higher ranked.

My analysis applies a different constraint-based framework. Inspired by Keller's (2000) work, I work with the idea of quantifiable constraint violation costs that are cumulated to determine the acceptability value of a sentence. A similar model is Harmonic Grammar (Legendre et al. 1990), which is implemented as a connectionist network. Both approaches are modifications (or a predecessor in the case of Harmonic Grammar) of Prince and Smolensky's (1993) standard OT in which all candidates but the winner are ungrammatical. Contrasting with this view, I rely on the idea that there are constraints whose violation leads to suboptimality without necessarily causing ungrammaticality. Such an approach is necessary in order to account for gradience and markedness, which neither standard OT nor stochastic OT can do, at least not without incorporating additional assumptions. Several studies have already presented constraint-based models that account for gradience. Müller (1999) proposed a modification of standard OT distinguishing between grammaticality constraints and markedness constraints: Some competitions would lead to more than one optimal candidate. In these cases, the optimal candidates would undergo a second competition, this time with a different set of constraints, namely markedness constraints. Coetzee (2004, 2006) assumes in his rank-ordering model of EVAL a critical cut-off that divides the constraint set into those constraints that a language is willing to violate and those that a language is not willing to violate. Violations of constraints below the cut-off point do not lead to ungrammaticality but to markedness. Gutiérrez-Bravo (2006) applies a similar approach in which he incorporates the notion of relative markedness into his analysis of word order variants in Spanish.

Keller (2000), also inspired by Uszkoreit's (1987) ideas on cumulativity in grammar, suggests that the cost of all violated constraints determines the final result (see also Pater 2009). He associates each constraint with a numerical weight representing the reduction in acceptability caused by a violation of this constraint. The acceptability of a structure is assumed to correspond to the weighted sum of the constraint violations it incurs. The weights correspond to constraint ranks (Keller 2000: 252/253). Cumulative numeric violation costs entail the notion of *ganging-up cumulativity*, if we understand cumulativity such that the effect of two weaker constraints might be stronger than the effect of one stronger constraint (see Jäger and Rosenbach 2006 for a discussion of different notions of cumulativity).

An approach based on specific violation costs and constraint cumulativity is particularly powerful when applied to experimental gradient acceptability data. It allows one to harvest the fruits of experimentally obtained gradient judgments, in order to propose theoretical insights on grammar.

2. Experimental approach

The word orders that are going to be studied with the experimental acceptability judgment test are shown in (38a) to (38h) for Catalan and (39a) to (39h) for Spanish. The context preceding this example is "no, no les claus" for Catalan and "no, no las llaves" for Spanish, meaning 'no, not the keys' (more details on the experimental material is given in Section 2.3). I am looking at four different positions of the c-focused object-DP (initial, after topic, in situ, and postposed). Furthermore, there is a variant with preverbal and one with postverbal subject for each of these four focus positions. The design thus contains the three independent variables focus position (A), subject position (B), and language (C), as in the table below. It does not separate object position and focus position (the interaction of these two factors would have to be dealt with in future research).

Table 1: Three-way ANOVA design

B: subject position	A: position of focused object							
	a_1: preposed/ sentence-initial		a_2: preposed/ after topic		a_3: in situ		a_4: postposed/ sentence-final	
	Cat.	Span.	Cat.	Span.	Cat.	Span.	Cat.	Span.
b_1: preverbal	(38a)	(39a)	(38c)	(39c)	(38e)	(39e)	(38g)	(39g)
b_2: postverbal	(38b)	(39b)	(38d)	(39d)	(38f)	(39f)	(38h)	(39h)

(38) a. *EL MÒBIL en Martí va perdre per casa.* $[O]_{c\text{-}F}SVP$ (Cat.)
the cell phone the Martí lost at home

b. *EL MÒBIL va perdre en Martí per casa.* $[O]_{c\text{-}F}VSP$ (Cat.)

c. *Per casa EL MÒBIL en Martí va perdre.* $P[O]_{c\text{-}F}SV$ (Cat.)

d. *Per casa EL MÒBIL va perdre en Martí.* $P[O]_{c\text{-}F}VS$ (Cat.)

e. *En Martí va perdre EL MÒBIL per casa.* $SV[O]_{c\text{-}F}P$ (Cat.)

f. *Va perdre en Martí EL MÒBIL per casa.* $VS[O]_{c\text{-}F}P$ (Cat.)

g. *En Martí va perdre per casa EL MÒBIL.* $SVP[O]_{c\text{-}F}$ (Cat.)

h. *Va perdre en Martí per casa EL MÒBIL.* $VSP[O]_{c\text{-}F}$ (Cat.)

(39)	a.	*EL MÓVIL Martín perdió por casa.*	$[O]_{c\text{-}F}$SVP	(Sp.)
	b.	*EL MÓVIL perdió Martín por casa.*	$[O]_{c\text{-}F}$VSP	(Sp.)
	c.	*Por casa EL MÓVIL Martín perdió.*	P$[O]_{c\text{-}F}$SV	(Sp.)
	d.	*Por casa EL MÓVIL perdió Martín.*	P$[O]_{c\text{-}F}$VS	(Sp.)
	e.	*Martín perdió EL MÓVIL por casa.*	SV$[O]_{c\text{-}F}$P	(Sp.)
	f.	*Perdió Martín EL MÓVIL por casa.*	VS$[O]_{c\text{-}F}$P	(Sp.)
	g.	*Martín perdió por casa EL MÓVIL.*	SVP$[O]_{c\text{-}F}$	(Sp.)
	h.	*Perdió Martín por casa EL MÓVIL.*	VSP$[O]_{c\text{-}F}$	(Sp.)

In the remainder of this section, I will present the method of data collection, the sampling plan, and the construction of the experimental test sentences.

2.1. Gradient acceptability test for auditory stimuli

Obtaining reliable and precise data concerning the focus-prosody interface constitutes a major methodological challenge because of the particular importance of intonation and context that have to be controlled (Zubizarreta 1999: 4218). Linguistic evidence, in particular grammaticality judgments, is far from always being solid (see, e.g., Schütze 1996), even without the additional challenges related to the focus-prosody interface. Data from spontaneous speech constitute undoubtedly the most natural source of evidence. While it is an appropriate source for determining differences in frequency between felicitous constructions, it is an inappropriate source for working on systematic differences in markedness because a number of (marked) constructions hardly occur or do not occur at all in spontaneous speech. The problem is that judgments cannot be reliable if (i) the suprasegmental characteristics of the phonetic realization and (ii) the focus-background structure are not carefully controlled for (Zubizarreta and Vergnaud 2005: Sec. 1). This was the motivation for developing in the scope of this study an acceptability judgment test for auditorily presented stimuli. In this test, subjects listen to a test sentence and give a nuanced judgment. They sat in a quiet room in front of a notebook computer with 15-inch display and wore closed stereo headphones.

Subjects rate the acceptability of the sentence on a *graphic rating* scale (also known under the label 'visual analogue scale'). Compared to classic categorical rating, graphic rating allows for the expression of fine nuances in a very intuitive way. This technique has been first described by Hayes and Patterson (1921). Validation tests based on different kinds of subjective perceptions or sensations report a discrimating capacity (or sensitivity) superior to the widespread categorical rating scale (Carlsson 1983; Reips and Funke 2008; Turner et al. 2008). In the present study, subjects draw a line on the computer screen by dragging a cross on a horizontal slider, which runs from 0 to 100. The longer

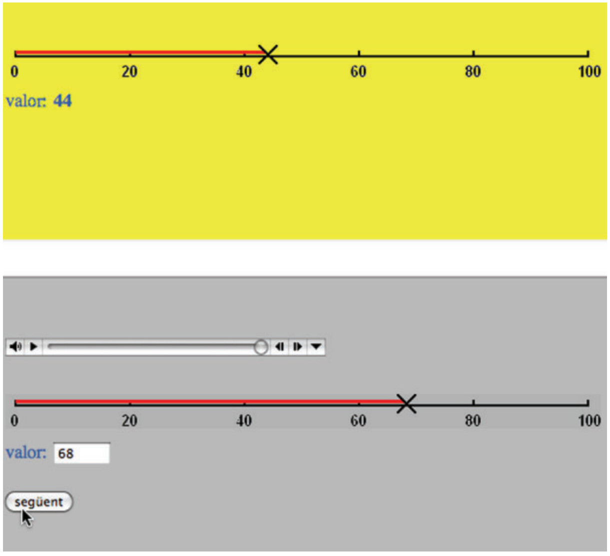

Figure 1: Acceptability judgment test

the line the higher is the degree of acceptability.[7] In addition, the actual value is displayed under the slider. Compared with paper and pencil versions of graphic rating tests (Adli 2005), the computer-based gradient acceptability test is easier to handle for subjects. The length of the line can be adjusted, and there is seamless integration with digital audio playback. The instrument is coded in the programming language *php* and runs in a normal browser window.

Upon clicking on the "next" button, playback of the following sentence starts. Subjects have the possibility of listening to the sentence again by clicking on the play button of the media toolbar; they can then adjust their judgment. However, once they click on the "next" button, they cannot go back. Subjects first judge an anchoring sentence. Its judgment remains visible in the upper part of the screen while they proceed through the test sentences (one by one, presented in a randomized order) in the lower part of the screen.

7. The graphic rating technique is preferred over magnitude estimation (Bard et al. 1996). Magnitude estimation presupposes a scale on which the intervals are not equidistant. I think that there is not sufficient empirical evidence for this assumption which could justify a deviation from the standard linear scale (see also Featherston 2009 on the issues of equidistance and the limits of an acceptability scale).

A suboptimal construction that mostly received a judgment in the intermediate scale range was chosen as the anchoring sentence.[8] An intermediate scale anchor generally helps to improve the precision of the instrument.

Subjects can determine not only the degree of acceptability by estimating the distance from the left (absolutely ungrammatical) and the right (obviously grammatical) endpoints, but can also compare it with a prior judgment to see how much their judgment differs.

Before starting the experimental phase, subjects sat down with the experimenter and were instructed and trained according to a nine-step procedure in which they essentially learned and practiced the concepts of *gradience* (as opposed to a binary perception of grammaticality), and *isolated grammaticality* (in order to minimize artifacts from irrelevant aspects such as pragmatic plausibility). The training started with written stimuli, namely sentences on paper strips, and continued with auditory stimuli at the computer. The sentences were presented in a randomized order. On average, the instructions lasted 15 minutes and the experimental phase 20 minutes. A demo version of the computer-based gradient acceptability judgment test for auditory stimuli as well as examples of the recordings of the experimental material are available online (the link can be found at *The Linguistic Review* on www.reference-global.com).

2.2. *Participants*

Data on Spanish and Catalan focus constructions was collected in February and March 2008 during fieldwork in Barcelona. The Spanish variety studied is the one spoken in Catalonia, which still belongs to the understudied varieties. A fundamental question that arises when working on the Spanish of Catalonia and on Catalan is how to deal with the fact that the majority of the inhabitants of Catalonia are bilingual. According to the 2001 census, 83.4 % of the persons between 20 and 39 years, and 90 % of the persons with high school degree or higher speak Catalan (IDESCAT 2007: 593). Basically all speakers of Catalan are also speakers of Spanish. If a speaker of Spanish is not a speaker of Catalan, she/he is in most cases a migrant from a monolingual Spanish-speaking region. There are two methodological options to deal with this situation: minimizing bilingualism (excluding it would not be possible), or embracing and

8. (i) is the anchoring sentence for the Catalan, and (ii) for the Spanish test version of the gradient acceptability judgment test.

(i)	??*Ahir*	*sí*	*que*	*a*	*la*	*muntanya*	*vam anar.*		(Cat.)
	yesterday	yes	that	to	the	mountains	went$_{1pl}$		
(ii)	??*Un*	*ratón*	*en*	*la*	*cocina*	*Marta*	*ayer*	*atrapó.*	(Sp.)
	a	mouse	in	the	kitchen	Marta	yesterday	caught	

AUTHOR'S COPY | AUTORENEXEMPLAR

systematically controlling it. The first option consists of a comparison of two independent samples, namely Spanish monolinguals from Catalonia and Catalan speakers with Catalan as their dominant language. One would look for the latter mainly outside of Barcelona, ideally in rural areas. This option was rejected for the following reasons: First, the samples would not be comparable, because the Catalan speakers would still be bilinguals (though the balance between the languages would be in favor of Catalan). Second, Spanish monolinguals of Catalonia often have a specific social trajectory, because they are usually migrants from other parts of Spain.[9] Consequently, the question would arise whether they should not rather count as speakers of the variety of their region (or country) of origin. Given that social difference and the issue of migration is not the object of this study, such a sampling would go against the aim of socially and dialectally homogeneous subjects.

I opted for the second option, because bilingualism represents the *natural* linguistic situation in Catalonia. This trend is even increasing (shown by the inverse correlation between knowledge of Catalan and age, IDESCAT 2007: 593). In this respect, the situation is not comparable to the one of second-generation migrants in essentially monolingual societies where bilingual speakers belong to a minority. Rather than trying (in vain) to exclude the phenomenon of bilingualism, I controlled for it, by selecting speakers with fairly balanced bilingualism, i.e. with a comparable proficiency of Spanish and Catalan. Thus, the design controls for this variable, but still respects the criterion of external validity.[10]

In bilingualism research a widely accepted view is the autonomy hypothesis according to which bilingual children develop separate grammatical L1 systems (e.g., Lindholm and Padilla 1978; Meisel 1986). Irrespective of this idea, there is a debate whether autonomous language systems influence each other or not, in other words whether the system of a bilingual person differs from the one of a monolingual person.

Some authors assume that possible interferences are restricted to the level of performance and that they do not affect competence (Paradis and Genesee 1996). Others, however, assume a systemic influence at the level of competence. There are two general positions on the conditions and the nature of influence. On the one hand, it has been argued that influence is driven by lan-

9. Since the Linguistic Normalization Act in 1983 a substantial proportion of the school education is conducted in Catalan (Pradilla 2001).
10. The control of bilingualism is also reflected in the instruction phase of the gradient acceptability test. At the onset, the language spoken between the subject and the equally bilingual experimenter was Catalan. Then the experimental phase for Catalan started. A short break followed during which the experimenter switched the language of conversation from Catalan to Spanish. Finally, the subjects continued with the experimental phase for Spanish.

guage dominance according to which structures of the dominant language can influence the structures of the non-dominant one (Yip and Matthews 2000; Bernardini and Schlyter 2004). On the other, it has been claimed that language dominance is not a relevant factor. Rather influence could be traced back to language-internal factors, mainly in terms of presence or absence of overlap between specific grammatical properties of the two languages (Hulk and Müller 2000). Kupisch (2007) argues that both dominance and language-internal factors play a role.

In this study, I adress the issue of potential dominance effects by selecting fairly balanced bilinguals (the notion of perfectly balanced bilinguals is probably too much of an idealization, see Grosjean 1982). One should bear in mind that the issue of interference has mainly been discussed for children during their acquisition, and authors who defend this position do not necessarily believe that interference persists beyond the main acquisition period (assuming that other factors such as attrition do not come into play). However, given those voices which claim that influence is not a mere performance phenomenon, and that it might occur independently of dominance, and that it might extend into adult age, I cannot exclude interferences between the grammatical systems of my bilingual subjects, precisely because the issue is still open and controversial. Importantly, if there is interference, I would *not* consider it as an artifact. Rather, it would reflect how the Spanish of Catalonia and Catalan *are*. The language contact situation is a constitutive element of these languages or varieties.[11]

The sample consisted of N = 54 carefully selected speakers of Catalan and the Spanish variety of Catalonia who can be considered as fairly balanced bilinguals.[12] A strict definition of bilingual language balance was applied during selection. Subjects actually had to grow up speaking both languages in their families (in most cases, they spoke Catalan with one parent and Spanish with the other parent or with a grandparent), they had to be active bilinguals by the age of six, and they had to have maintained an active daily use of both languages. Nevertheless, if the author and the two local assistants (who themselves conformed to these criteria) suspected any language dominance, the subject was excluded. Furthermore, subjects self-evaluated, in a final questionnaire, their level of bilingualism on a 7-point semantic differential scale, with the

11. The same argument would also hold for a number of other multilingual regions in the world (many Western and Central African countries, various Spanish-speaking South-American countries with substantial proportion of indigenous population, etc.).
12. As a side note: A design with balanced bilinguals has an advantage over a design with two monolingual samples in terms of statistical power. Instead of comparing two independent samples of monolinguals consisting of N/2 subjects each, one compares two dependent samples of N subjects each. It thus allows more reliable conclusions, which is important when we want to look at nuanced differences of markedness.

endpoints "only Catalan" (1) and "only Spanish" (7). The mean value 4.1 corresponds to the answer "totally balanced" (4), confirming the adequacy of the selection criteria.

The sample consisted of 60 % women and 40 % men. Their age ranged from 17 to 48 (the mean age was 27 and a half). Dialectal variation was nearly absent (94 % grew up with the Central Catalan dialect). The data of 51 fulfilled the validity criteria and were used in this analysis (the bilingualism of two people did not seem sufficiently balanced, and one participant was excluded due to insufficient cooperation).

2.3. Experimental material, test constructions, and mean values

An important methodological issue concerns the question whether specific phonetic characteristics should be superposed on the test material (e.g., reading exactly according to a predefined tone structure), or whether natural speech behavior should be given priority. Face (2002: 90) has shown in his study of phonetic and phonological properties of c-focus in Spanish that there are multiple intonational strategies: (i) a L*+H pitch accent with a higher F^0 peak compared to declaratives with wide focus, (ii) a focal L+H* pitch accent, (iii) H^-, or (iv) L^- following the c-focused word. As for Catalan, Prieto (2002a: 414, 2002b: Sec. 4.5.2) illustrates a pitch contour of c-focus with a high peak on the c-focused word followed by a low, flat F^0. She describes the contrastive pitch accent as H*+L (however, she also points out that there is more than one intonational realization). Estebas-Vilaplana (2003) claims that c-focus is realized as H^*L^-, i.e., that the L corresponds to a phrase accent (see also Estebas-Vilaplana 2000).

However, it is not clear which factors determine the choice of these intonational variants. Furthermore, suprasegmental features are difficult to identify, especially in complex constructions, given that they do not have unique phonetic correlates (Zubizarreta and Vergnaud 2005: Sec. 1). A more appropriate approach seems to consist of recording linguistically trained native speakers. Since they are familiar with marked and ungrammatical constructions, they can utter them more easily in a fluid and natural manner. Linguistic training also ensures full understanding of the concepts of contrastive focus, presupposition, emphasis, context, etc. For this purpose, one male and one female graduate student of linguistics from Barcelona, both balanced bilinguals, were recorded. They were supposed to read the test sentences in a natural way, i.e., without trying to imitate a predetermined pitch contour and without producing exaggerated 'laboratory speech'. The material is preceded by a context to make the presupposition, in order to elicit an unambiguous interpretation as contrastive stress and not as emphatic stress (examples can be found in Zubizarreta 1999: 4229). The context to c-focus is salient due to the negation of the presupposition.

Different lexical material for verbs and nouns is used in each sentence in order to avoid artifacts due to repetition (such as habituation or priming effects).[13] The material was digitally recorded using the ProTools system and Sony ECM-77B condenser microphones. Each item was recorded several times, and the

13. (i) to (viii) represent one of the two sets of lexical variants of the original Catalan test sentences. (ix) to (xvi) represent one of the two sets of lexical variants of the Spanish test sentences.

(i)	*No, no el disc.* *EL LLIBRE en Joan va canviar a la* FNAC.	(Cat.)
(ii)	*No, no les claus.* *EL MÒBIL va perdre en Martí per casa.*	(Cat.)
(iii)	*No, no els diners.* *A la mansió LES JOIES el lladre va robar.*	(Cat.)
(iv)	*No, no el pont.* *A Madrid L'EDIFICI va construir l'enginyer.*	(Cat.)
(v)	*No, no la paret.* *Picasso va pintar EL QUADRE a París.*	(Cat.)
(vi)	*No, no la taula.* *Va arreglar el fuster LA CADIRA al taller.*	(Cat.)
(vii)	*No, no els exàmens.* *El professor va corregir a l' escola ELS TREBALLS.*	(Cat.)
(viii)	*No, no un crani.* *Va descobrir l' arqueòleg a Egipte LA MÒMIA.*	(Cat.)
(ix)	*No, no la puerta.* *LA VENTANA Juan cerró en la casa.*	(Sp.)
(x)	*No, no la verdura.* *EL CHULETÓN comió Miguel en el restaurante.*	(Sp.)
(xi)	*No, no la carta.* *En la oficina EL PAQUETE María envió.*	(Sp.)
(xii)	*No, no el Burdeos.* *En el restaurante EL RIOJA bebió Carmen.*	(Sp.)
(xiii)	*No, no los pantalones.* *Laura compró LA CAMISA en Zara.*	(Sp.)
(xiv)	*No, no matemáticas.* *Estudió Mercedes SOCIALES en la escuela.*	(Sp.)
(xv)	*No, no el bolero.* *Ana cantó en el coro LA TRAVIATA.*	(Sp.)
(xvi)	*No, no la guitarra.* *Tocó Pedro en la clase EL PIANO.*	(Sp.)

best realization (most natural reading, with clearly perceivable but not exaggerated nuclear stress according to the opinion of three persons) was chosen for the final experimental material. The male speaker recorded approximately one half and the female speaker the other half of the final material.

Each construction was presented in two lexical variants to every participant. The dependent variable is the mean value of the two lexical variants (statistical reliability, in terms of Cronbach's α and the intercorrelation coefficient, is generally higher if the result for a specific construction is based on more than one item or lexical variant, Adli 2005). Subjects saw $2 \times 8 = 16$ Catalan experimental sentences plus one anchoring sentence, $2 \times 8 = 16$ Spanish experimental test sentences plus one anchoring sentence, to which one has to add 10 sentences from the instruction and training phase, which is a total of 44 gradient judgments per person.[14]

The gradient judgment data is analyzed with a three-way analysis of variance, i.e., a design with three independent variables as given in Table 1 (position of the c-focused object, subject position, and language). The following figures show the mean gradient acceptability values for each construction. They directly reflect the answers of the subjects on the graphic rating scale from 0 to 100. A continuous line connects the results for constructions with preverbal subject, and a dashed line the results for constructions with postverbal subject. Furthermore, lines with square markers correspond to Catalan sentences, while lines with diamond-shaped markers correspond to Spanish sentences. The four horizontal categories represent the different positions of the c-focused object. In view of a better readability, Figure 3 and Figure 4 show the results for Catalan and Spanish separately.[15]

The fact that even the best candidates of each set, (38e) and (39e), do not obtain a mean value close to the scale maximum 100 but a value below 75, does not mean that these constructions are suboptimal. Rather, it reflects a common

---

14. In addition, the experimental session included acceptability judgments on other linguistic phenomena such as interrogatives, which are not the topic of the present work. Although those stimuli are not filler sentences in the narrow sense, they nevertheless contribute to a diversification of the material.
15. The following table lists mean value and standard deviation for all sentences:

	(38a)	(38b)	(38c)	(38d)	(38e)	(38f)	(38g)	(38h)	(39a)	(39b)	(39c)	(39d)	(39e)	(39f)	(39h)	(39g)
$\bar{x}$	65.8	59.3	38.6	45.2	69.2	60.3	59.7	46.7	53.8	61.8	41.8	52.7	74.6	55.6	69.8	44.0
s	20.2	20.6	14.9	17.4	20.9	18.4	20.4	18.8	19.3	18.3	17.0	19.2	21.0	18.4	20.0	20.3

All standard deviations s fall in the range between approximately 15 and 21. In line with observations from previous studies (see Adli 2005), judgments on suboptimal constructions – those with the lowest acceptability value – have a higher degree of consistency, i.e., their standard deviation is lower. In particular, we find $s = 14.9$ for (38c), $s = 17.0$ with (39c). This contrasts clearly with sentences that come with high acceptability values: for example, $s = 20.9$ for (38e), $s = 21.0$ for (39e).

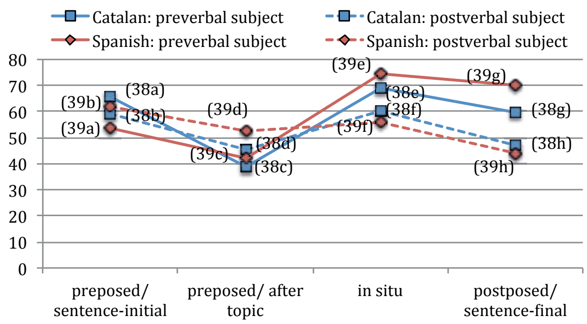

Figure 2: Acceptability judgments for Catalan and Spanish

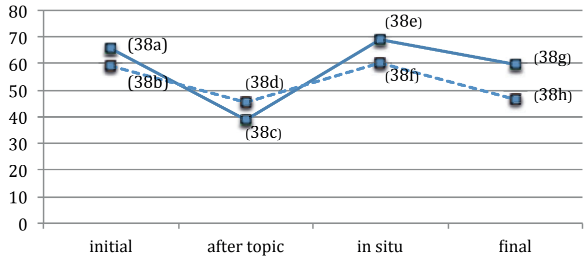

Figure 3: Judgments for Catalan

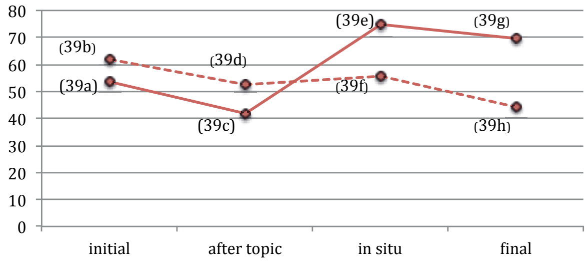

Figure 4: Judgments for Spanish

psychometrical phenomenon, the central tendency of judgment first described by Hollingworth (1910), according to which most subjects avoid the lower and higher end of the scale (see also Crawford et al. 2000; Albaum et al. 2007). In other words, the really operative scale range is smaller, at least in terms of arithmetic means.[16] An important consequence for the present study is that we can consider the acceptability value of the best candidate of each set as a near-to-maximum value (more on this in Section 4).

2.4. Analysis of variance

The gradient judgment data is statistically analyzed with a multi-way analysis of variance for repeated measures using the software SPSS 16. Analysis of variance distinguishes between the effects of various independent variables, and their interactions (Lindman 1974). Each single measurement x_{ijm} is decomposed into several systematic effects and an error value (cf. Bortz and Schuster 2010: Ch. 12). (40), for example, shows the structural components of a two-way analysis of variance with between-subject variables A and B:

(40) $$x_{ijm} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijm}$$

x_{ijm} is the single measurement of subject *m* at the level combination *ij*. μ is the mean value of all measurements (i.e., it is the study-specific general measurement level) and thus constant for all single measurements x_{ijm}. α_i is the specific effect of level *i* of variable A (e.g., α_3 is the in-situ focus position in Table 1), β_j is the specific effect of level *j* of variable B (e.g., β_1 is the preverbal subject position), $\alpha\beta_{ij}$ is the interaction effect of the variable level combination ab_{ij}, and ε_{ijm} is the measurement error of the single measurement. The formula illustrates that main and interaction effects are independent. If *main effect* A is significant, it means that at least one focus position is different from the others (technically speaking, the null hypothesis $\alpha_1 = \alpha_2 = \cdots = \alpha_i = \cdots = \alpha_p = 0$ is then rejected). An *interaction effect* A $\times$ B means that there is an effect of the combination of focus position and subject position that cannot be attributed to a sole main effect (e.g., the effect of focus position is substantially different for preverbal than for postverbal subjects). Other main and interaction effect tests are analogous.

At several points *simple main effect* tests are calculated to gain a more detailed view. The simple main effect test $A|b_j$ tells us whether main effect A has an effect, if only level *j* of the independent variable B is taken into consideration (e.g., $A|b_1$ tells us whether there is a difference between the four focus positions looking only at constructions with preverbal subject). Likewise, a_n

16. As with all normal distributions, the sample also contains a few cases who give substantially lower or higher ratings than the mean value. The maximal rating of 100 has been assigned by two subjects to sentence (38e) and by one subject to (39e). (38e) and (39e) are the constructions with the highest mean value. Likewise, (38c) and (39c) are the constructions with the lowest mean value. No subject rated them with 100.

vs. $a_m|b_j$ tells us whether the contrast between two levels of A has an effect under b_j (e.g., a_3 vs. $a_4|b_2$ tells us whether there is a difference between the in situ and the final focus position for constructions with postverbal subject). The size of the effect can be small or big. This information is provided by the *partial* η^2 value, which is a supplementary descriptive measure (Cohen 1973). It is more informative than the difference between the mean values $\Delta\overline{X}$, because the deviation or dispersion of the data is taken into consideration (however, $\Delta\overline{X}$ is often indicated since it is an intuitive value reflecting the size of the mean rating difference on the scale from 0 to 100). The partial η^2 is a commonly used measure in psychology (its systematic report is even required by the American Psychological Association 2001: 25), but to my knowledge only very few linguistic studies (e.g., Kondo-Brown 2005) have included it.[17]

3. Analyzing gradience

The statistical analysis of the gradient judgments can be used to draw a picture of the differences in markedness and the presumed importance of a constraint. However, even though the judgments are fine-grained, it is not sufficient to determine exact costs within a model of cumulative constraint violation. Such a model is too complex in order to be able to grasp a result with the naked eye. If a sentence scores very high, we know that none of the constraints it violates has a high violation cost. However, if a sentences scores low, we only know that the sum of all violation costs of the constraints it violates is high. But we do not know whether most of the total violation cost goes back to a single important constraint (while the others are less decisive), whether it goes back to two equally important constraints, or whether all violated constraints contribute more or less equally to the result. Some algebraic approach is necessary for dealing with this issue to which I come in the next section. First, we identify the

17. The different hypotheses are calculated at $\alpha = \beta$, i.e., balanced error probabilities or "fair hypothesis testing" (Erdfelder and Bredenkamp 1994; Erdfelder 2010) with medium effect size. α is the probability with which the null hypothesis is erroneously rejected. The value β is the probability with which the alternative hypothesis is erroneously rejected. Fair testing means that both conclusions are drawn on the same error probability: the acceptability of the sentences is identical (i.e., a non-significant result), or the acceptability differs (i.e., a significant result). The effect size ε is the minimal difference that has to exist in order to achieve statistical significance. Controlling for ε is important to avoid the situation that minimal, irrelevant differences lead to significance (ε too small), or that really large differences are still not detected by the test (ε too large). Medium ε corresponds to the usual convention according to Cohen (1988). The values for α and β at medium ε are calculated with the software G*Power 3 (Faul et al. 2007): We obtain for most tests $\alpha = \beta = 5.6\,\%$, except for main effect A ('focus position') for which we obtain $\alpha = \beta = 1.9\,\%$. These are small error probabilities at decent test sensitivity. The significance tests are based as usual on the F-statistics.

constraints based on a discussion of the structural properties of the sentences and try to come up with a plausible hypothesis on constraint ranking.

I start by looking at the main and interaction effects in order to obtain a general view of the data. Then, the details will be analyzed in the following subsections. The results of the analysis of variance reveal that main effect A 'focus position' ($F(3,150) = 57.085$, $p < 0.000$) and main effect B 'subject position' ($F(1,50) = 68.499$, $p < 0.000$) are significant. In other words, there is a difference between the different positions of the c-focused object and also a difference between preverbal and postverbal subjects. The significant interaction $A \times B$ ($F(3,150) = 36.658$, $p < 0.000$) shows that there is also an effect of focus position, which is specific to the subject position. These main effects are also detected when we look at both languages separately. Focus position, subject position, and their interaction is significant for Catalan as well as for Spanish (all simple main effects $A|c_1$, $A|c_2$, $B|c_1$, $B|c_2$, $A \times B|c_1$, $A \times B|c_2$: $p < 0.000$).[18]

The contrasts between Catalan and the Spanish variety of Catalonia (henceforth simply called "Spanish") come to fore in the interaction effects. We observe a language-specific effect of the focus position in the significant interaction $A \times C$ ($F(3,150) = 7.561$, $p < 0.000$). Not only the focus position, but also even the interaction between focus position and subject position is language-specific, as is revealed by the triple interaction $A \times B \times C$ ($F(3,150) = 13.973$, $p < 0.000$). Interestingly, the interaction between focus position and subject position is much more pronounced in Spanish (partial $\eta^2_{A\times B|c2} = 0.478$) than in Catalan (partial $\eta^2_{A\times B|c1} = 0.178$). We will come back to this point later.

3.1. *In situ and sentence-final c-focus*

I define the constraint FOCP$_{\text{C-F}}$ as in (41) which assumes that a fronted c-focus is under CP, more precisely it moves to FocP.[19]

(41) FOCP$_{\text{C-F}}$: [XP … $_{\text{c-F}}$X/Y] in FocP

(41) is violated by the sentences with in situ and final c-focus, i.e., (38e) to (38h) in Catalan and (39e) to (39h) in Spanish. Some of these constructions

18. Main effect C 'language' is not significant ($F(1,50) = 0.437$, $p < 0.000$). This result is expected (if the instrument works reliably and the material correctly constructed) since there is no a-priori difference between the Catalan candidate set (38a) to (38h) and the Spanish one (39a) to (39h). What is different is the markedness ranking within each set.
19. (41) is a modification of Gabriel's (2007: 244) SPEC,TP/FOC constraint in (i) according to which the fronted focus moves to TP.

(i) [XP … $_{\text{[Foc]}}$X/Y] in Spec,TP.

score very high, in particular (38e) and (39e). In both Catalan and Spanish, constructions with in situ c-focused object (and preverbal subject) have the highest acceptability values. This result indicates that any constraint that requires the c-focused object not to remain in situ must be ranked very low. Consequently FOCP$_{\text{C-F}}$ in (41) which is violated when the c-focus is not fronted as well as ALIGNFOC in (34), which is violated when the focus constituent does not right-align with the I-phrase, are ranked very low, i.e., they have a low violation cost. The fact that some of the constructions (38e) to (38h) and (39e) to (39h) score very low, e.g., (38h) and (39h), is due to the cumulative effect, i.e., mainly due to the violation of other, presumably higher-rated constraints.

Furthermore, a sentence-final position of the c-focused object is only slightly more marked (again, looking only at preverbal subjects): The difference between the sentence-final and the in situ position is significant, but the effect sizes are rather modest. For the Catalan constructions we observe (38e) > (38g) with a partial η^2 of 0.192 (a_3 vs. $a_4|b_1c_1$: mean difference $\Delta\overline{X} = 9.5$, $p < 0.001$) and for Spanish (39e) > (39g) with a partial η^2 of 0.074 (a_3 vs. $a_4|b_1c_2$: $\Delta\overline{X} = 4.8$, $p < 0.047$). The difference is particularly small for Spanish, where sentence-final c-focus is nearly (but not fully) on a par with in situ c-focus. Since we assume with Zubizarreta (1998) that sentence-final focus is the result of *p*-movement of the PP-adjunct, this result indicates that the Stay-ϕ constraint in (31b) is also ranked low, but presumably not as low as ALIGNFOC and FOCP$_{\text{C-F}}$.

Another indication that supports the assumption that in-situ position (with preverbal subject) is the most unmarked option for contrastive focus comes from multiple focus constructions. It is possible to have multiple in-situ c-focus, but it is not possible to have multiple preposed c-focus, as the Spanish examples (42a) and (42b) show (Zubizarreta 1999: 4241).

(42) a. *Estoy segura de que Adán le dio [la* (Sp.)
am$_{\text{1Sg}}$ sure about that Adam her gave the
MANZANA]$_{\text{c-F}}$ [*a EVA*]$_{\text{c-F}}$ *(y la PERA a MARÍA).*
apple to Eve and the pear to Maria

b. **Estoy segura de que* [*la MANZANA*]$_{\text{c-F}}$, [*a* (Sp.)
am$_{\text{1Sg}}$ sure about that the apple to
EVA]$_{\text{c-F}}$ *le dio Adán.*
Eve her gave Adam

Interestingly, Gabriel's (2007: 285) results based on elicited Spanish c-focus constructions show that (i) fronted focus never occurs in his transcripts, and (ii) the in situ position is the most common form (though the speakers prefer not to have further material following the direct object, i.e., they prefer in situ c-focus at the right edge of the I-phrase).

AUTHOR'S COPY | AUTORENEXEMPLAR

In Catalan, the preverbal focus position (with preverbal subject) (38a) has a degree of acceptability as high as the one with in situ focus, and it still ranges before the sentence-final focus position: (38e) = (38a), and (38a) > (38g). I had suggested further above the constraint *COPY$_{[+F]}$ in (32b), which is violated when [+F]-marked material is displaced. The high rating of (38a), which violates *COPY$_{[+F]}$, shows that this constraint must also be ranked low in Catalan. This is further supported by the fact that sentence-initial and in situ c-focus constructions with postverbal subject, (38b) and (38f), also show identical levels of grammaticality (a_1 vs. $a_3|b_2c_1$: $\Delta\overline{X} = 1.0$; $p < 0.746$). In Spanish, *COPY$_{[+F]}$ must be ranked relatively low, too, but not as low as in Catalan because the constructions with preverbal c-focus (and preverbal subject) are more marked than the constructions with in situ or sentence-final c-focus. In other words, (39e) and (39g) have higher acceptability values than (39a) and (39c).

We already know that there is some constraint requiring focus to be sentence-initial, presumably ranked high in Catalan. However, in order to understand the complete picture in both languages, we need to turn our attention to the role of the subject position.

3.2. Preverbal and postverbal subjects

Preposed focus in Spanish requires the subject to be postverbal (Zubizarreta 1999: 4241) . We see this effect with sentence-initial c-focus, (39b) > (39a) ($B|a_1c_2$: $\Delta\overline{X} = 8.1$, partial $\eta^2 = 0.149$, $F(1,52) = 9.131$, $p < 0.004$) and even somewhat stronger with the topic-focus order, (39d) > (39c) ($B|a_2c_2$: $\Delta\overline{X} = 10.9$, partial $\eta^2 = 0.408$, $F(1,52) = 35.861$, $p < 0.000$).

However, the *interaction* between focus position and subject position, already mentioned above, has gone unnoticed in the literature. Zubizarreta (1999: 4236) argues that both SVOP and VSOP are well-formed with c-focus on the object, as her judgments in the examples (43a) and (43b) show. She assumes that there is flexibility in placing nuclear stress either on the object DP or on the PP.

(43) a. *Ayer colgaron los alumnos de la* (Sp.)
yesterday attached the pupils of the
primaria la bandera FRANCESA en el
elementary school the flag FRENCH to the
mástil (no la INGLESA).
mast (not the BRITISH)

b. *El gato con botas escondió el QUESO* (Sp.)
the cat with boots hid the cheese

debajo de	*la*	*cama*	(*no*	*el*	*BOLLO*).
under	the	bed	(not	the	PASTRY)

While the judgment for the SV[O]$_{\text{c-F}}$P order in (43b) with in situ c-focus and preverbal subject is undisputed, the VS[O]$_{\text{c-F}}$P order with postverbal subject in (43a) can only be considered as highly marked. The experimental results are clear: compared to (39e) with preverbal subject, the variant of the in situ c-focus construction with postverbal subject (39f) comes along with a pronounced decrease in grammaticality ($B|a_3c_2$: $\Delta\overline{X} = 19.0$, partial $\eta^2 = 0.465$, $F(1,52) = 45.213$, $p < 0.000$). The decrease due to the postverbal subject is even larger with sentence-final c-focus, as can be seen in the difference between (39g) and (39h) ($B|a_4c_2$: $\Delta\overline{X} = 25.8$, partial $\eta^2 = 0.627$, $F(1,52) = 87.265$, $p < 0.000$). The interaction between focus position and subject position in Spanish is expressed by the following principle (recall that $P > Q$ means P is preferred over Q).

(44) If the contrastivly focused object is preposed, then postverbal subjects are preferred: [O]$_{\text{c-F}}$VS > [O]$_{\text{c-F}}$SV.
If the contrastively focused object is in situ or postposed, then preverbal subjects are preferred: SV[O]$_{\text{c-F}}$ > VS[O]$_{\text{c-F}}$.

(44) does not hold for Catalan. Preverbal subjects are preferred with c-focused objects at sentence-initial position. We observe (38a) > (38b) ($B|a_1c_1$: $\Delta\overline{X} = 6.5$, partial $\eta^2 = 0.118$, $F(1,50) = 6.688$, $p < 0.013$). Only the (suboptimal) topic-focus order shows that (38d) with postverbal subject is somewhat less marked than (38c) with preverbal subject ($B|a_2c_1$: $\Delta\overline{X} = 6.6$, partial $\eta^2 = 0.158$, $F(1,50) = 9.356$, $p < 0.004$). We also notice that the contrast between preverbal and postverbal subjects in constructions with in situ or postposed c-focus is much less pronounced than in Spanish, as the difference between (38e) and (38f) and also between (38g) and (38h) shows. I therefore qualify the Catalan VSO order in (38f) as somewhat suboptimal, but not as ungrammatical, as does López (2009: Ch. 5.2.5).

An interesting observation is that in Spanish the preverbal focus position is slightly marked, while in Catalan this focus position is essentially as acceptable as the in situ position. If we choose among the Spanish constructions with preverbal focus the one with the highest acceptability value, namely (39b), and compare it with the fully unmarked in situ form (39e), then we observe a notably lower degree of acceptability of (39b) ($\Delta\overline{X} = 12.7$, partial $\eta^2 = 0.311$, $p < 0.000$). However, the same procedure for Catalan, i.e., the comparison of (38a) with the fully unmarked in situ form (38e), reveals an identical level of acceptability ($\Delta\overline{X} = 3.4$, partial $\eta^2 = 0.033$, $p < 0.198$). We can therefore say that in Catalan both *foco antepuesto* and in situ placement are the fully unmarked forms and optional variants in terms of identical acceptability. In

Spanish, *foco antepuesto*, though part of the repertory, is somewhat marked.

These facts now have to be translated into constraints. We need to assume different constraints in order to explain that (i) postverbal subjects with in situ or postposed focus are more marked both for Catalan and Spanish, (ii) preverbal subjects are generally more marked with preposed focus in Spanish and more marked with the focus-topic order in Catalan. In a model based on the cumulation of violation costs, (severe) markedness of the Spanish verb-initial constructions (39f) and (39h), and (slight) markedness of the Catalan pendants (38f) and (38h) cannot be accounted for solely by violation of FULLINT, repeated in (45). One also has to include the violation of some other constraint. The FULLINT constraint is always violated when there is a postverbal subject (contrasting with Gabriel 2007 I do not assume that a c-focused XP avoids its violation). We have seen that the variant with postverbal subject is preferred in certain constructions. FULLINT must therefore be ranked rather low in Catalan and very low in Spanish.

Gabriel (2007: 244/245) proposed the constraint in (46) to explain why constructions with a *v*P-internal subject followed by an argument such as (47) are highly marked.

(45) FULLINT:
Parse lexical conceptual structure. Failed by expletives and auxiliary do.

(46) *[VP SUBJ . . . ARG]

(47) **Compró María un diario.* (Sp.)
bought Maria a newspaper

The constructions (38f), (38h), (39f), and (39h) violate – in addition to FULLINT – the constraint *[VPSUBJ . . . ARG] in (46) because the *v*P-internal subject is followed by the focused object. A closer look shows that *[VP SUBJ . . . ARG] cannot be the missing piece for explaining the pattern with the constructions with a postverbal subject: (46) would imply that the sentences with in situ and sentence-final focus have lower acceptability values than the ones with preposed focus. However, we observe that (39d) is not better than (39f) in Spanish (a_2 vs. $a_3|b_2c_2$: $\Delta\overline{X} = 2.9$, partial $\eta^2 = 0.045$, $p < 0.123$) and that (38d) is even worse than (38f) in Catalan (a_2 vs. $a_3|b_2c_1$: $\Delta\overline{X} = 15.1$, partial $\eta^2 = 0.509$, $p < 0.000$). Rather we need a constraint that explains the high degree of markedness of constructions with postverbal subject and sentence-final focus such as (38h) and (39h) in order to account for the data. I suggest the constraint given in (48), which is violated by constructions with postverbal subjects, in which the subject dominates but not immediately dominates a c-focused argument due to the interposition of a *p*-moved constituent. This constraint explains the difference in markedness both in Catalan (38f) > (38h)

(a_3 vs. $a_4|b_2c_1$: $\Delta\overline{X} = 13.6$, partial $\eta^2 = 0.407$, $p < 0.000$) and in Spanish (39f) > (39h) (a_3 vs. $a_4|b_2c_2$: $\Delta\overline{X} = 11.6$, partial $\eta^2 = 0.328$, $p < 0.000$). The relevant facts are repeated, with (49a) > (49b).

(48) *[$_{vP}$ SUBJ ϕ FOC]

(49) a. *Perdió Martín EL MÓVIL en casa.* (Sp.)
lost Martín the cell phone at home
b. *Perdió Martín en casa EL MÓVIL.* (Sp.)

Finally, we need to explain the preference for postverbal subjects in preposed focus constructions, which applies to any preposed focus position in Spanish but is limited to the topic-focus order in Catalan. The Spanish facts can be explained with a constraint that I call CP-VS. A first formulation is given in (50).

(50) CP-VS:
Whenever there is a CP layer, V^0 dominates the subject.

In Catalan a postverbal subject is only preferred with the topic-focus order. I suggest that the reason is due to the fact that the CP is larger in the topic-focus order (38c) and (38d), because it contains both a topic and a focus projection. Not only is (38d) less marked than (38c) (TopP and FocP), the following judgments (not experimentally tested, though confirmed by several native speakers) corroborate the assumption that in Catalan postverbal subject is preferred with larger CPs: (51b) is less marked than (51a) (FocP and TopP), and (52b) is less marked than (52a) (COMP and FocP).

(51) a. ??*EL MÒBIL per casa en Martí va perdre.* (Cat.)
the cell phone at home the Martín lost
b. ?*EL MÒBIL per casa va perdre en Martí.* (Cat.)

(52) a. ??*Penso que EL MÒBIL en Martí va perdre.* (Cat.)
think$_{1sg}$ that the cell phone the Martín lost
b. ?*Penso que EL MÒBIL va perdre en Martí.* (Cat.)

I therefore propose a generalization as in (53). This constraint must be ranked relatively high because of the high degree of markedness of Catalan (38c) and Spanish (39c).

(53) CP-VS:
Whenever there is a CP layer with n or more projections, V^0 dominates the subject ($n = 1$ for Spanish, $n = 2$ for Catalan).

3.3. Focus position at the left periphery

The experimental results have revealed that preposed focus is preferred in sentence-initial position (a_1 vs. a_2: partial $\eta^2 = 0.637$, $F(1,50) = 87.722$, $p < 0.000$). The results for both Catalan and Spanish are consistent: We observe this preference with preverbal subjects in Catalan (38a) > (38c) (a_1 vs. $a_2|b_1c_1$: $\Delta\overline{X} = 27.2$, partial $\eta^2 = 0.662$, $p < 0.000$), with postverbal subjects in Catalan (38b) > (38d) (a_1 vs. $a_2|b_2c_1$: $\Delta\overline{X} = 14.1$, partial $\eta^2 = 0.301$, $p < 0.000$), with preverbal subjects in Spanish (39a) > (39c) (a_1 vs. $a_2|b_1c_2$: $\Delta\overline{X} = 12.4$, partial $\eta^2 = 0.414$, $p < 0.000$), and with postverbal subjects in Spanish (39b) > (39d) (simple main effect a_1 vs. $a_2|b_2c_2$: $\Delta\overline{X} = 9.4$, partial $\eta^2 = 0.210$, $p < 0.001$).

These results contradict Zubizarreta's (1998: 118, 1999: 4241) view who assumes that preposed focus in Spanish must be (left-)adjacent to the verb, stating: "The following orders are not an option: *Emphatic-XP-V…, *Focus-XP-V…". She assumes that this adjacency requirement is the reason why subjects have to be postverbal. Our data show that there is no adjacency requirement, at least in the Spanish variety of Catalonia. We have seen that the P[O]$_{\text{c-F}}$VS order (39d) is not any better than the [O]$_{\text{c-F}}$SVP order (39a) ($\Delta\overline{X} = 1.1$, partial $\eta^2 = 0.005$, $p < 0.628$) – in Catalan (38d) is even worse than (38a) ($\Delta\overline{X} = 20.6$, partial $\eta^2 = 0.55$, $p < 0.000$). Rather, there seems to be a principle that directly effects subject position.

The point that preposed focus must be adjacent to the verb and follow a topic (Vallduví 1992) was one of the main arguments against the full adoption of Rizzi's (1997) CP system for Spanish. With respect to topics, the position under CP is less controversial; multiple topics are possible and they can appear in a different order (Campos and Zampini 1990; Goodall 2001). However, starting with Zubizarreta (1998: 122/123), several authors have taken the stance that preposed focus lands at the Spec of TP (or Pole) (Barbosa 2001: 33; Gabriel 2007). The experimental results indicate that fronted topic and focus phrases occupy a position under CP, both in Catalan (also claimed by Bonet and Solà 1986) and in the Spanish variety of Catalonia.

Independent evidence comes from the results of a different study from the author, in which I analyzed the two orders *wh*-before-adverb such as (54a) or (54b) and adverb-before-*wh* such as (55a) or (55b) in interrogatives of the Spanish variety of Catalonia with a gradient grammaticality judgment test for written stimuli. The results showed a clear preference for the *wh*-before-adverb order, i.e., (54a) and (54b) are preferred over (55a) and (55b).

(54) a. *A quién anoche Federico le pidió la mano?* (Sp.)
of whom last night Fred CL asked for hand in marriage

b. *Por qué hoy Lola ha pedido el finiquito?* (Sp.)
why today Lola has asked the receipt

(55) a. *El domingo a quién Esperanza le* (Sp.)
the Sunday to whom Esperanza CL
confesó el pecado?
confessed the sin

b. *Anoche por qué Pablo contó el secreto?* (Sp.)
last night why Paul told the secret

Assuming that initial temporal adverbs move from an IP-adjoined position (Solà 2002; or VP-right-adjoined position according to Gutiérrez-Bravo to appear: 26) to a topic position (Suñer 1994; Gutiérrez-Bravo 2006: 149/150,191), which is TopP in a Split-CP account, this finding contradicts the view according to which topics typically precede wh-phrases in root clauses (Meyer 1972; López 1999; Gutiérrez-Bravo 2006). This point had been partly derived from the assumption, set forth by Goodall (1991), that fronted *wh*-arguments in Spanish questions occupy [Spec, T] ([Spec, I], in his terms). However, the definition in (1) states the correspondence between a focused element and a *wh*-element (see also Erteschik-Shir 1986). Therefore it seems safe to conclude that sentence-initial topic followed by focus or *wh* is dispreferred, both in the Spanish variety of Catalonia and in Catalan. Recall Rizzi's (1997: 296) articulation of the topic-focus system in (28) above. The results suggest that the preferred placement of the topic projections is below focus. It is possible above, though more marked.

I assume the constraint CP-LEFTFOC defined as in (56), which is violated by the order topic-focus, but not by focus-topic. CP-LEFTFOC is a sort of counterpart to ALIGNFOC.

(56) CP-LEFTFOC:
Whenever there is a CP layer, the focus constituent left-aligns with the I-phrase.

CP-LEFTFOC also explains the preference for initial *wh* in constituent questions as in (54a) and (54b). It must obviously be higher ranked than Costa's (1998) TOPICFIRST constraint, reformulated by Gutiérrez-Bravo (2007: 121) as in (57) to allow multiple violations. As a matter of fact, TOPICFIRST is not relevant for explaining the experimental results for Catalan and Spanish.

(57) TOPICFIRST:
Topics are sentence initial; nontopics cannot be topicalized. Violated once by every head X^0 or every maximal projection XP in the left periphery that c-commands a topic.

I will explain the judgment pattern in both Catalan and Spanish with the following constraints: CP-VS, CP-LEFTFOC, *[VPSUBJϕFOC], FULLINT, STAY-ϕ, *COPY$_{[+F]}$, FOCP$_{\text{C-F}}$, and FOCUSSALIENCY (the latter only applies to Spanish and will be discussed in Section 4.7). ALIGNFOC, *[VPSUBJ ... ARG], and TOPICFIRST will not be necessary to account for the data.

3.4. Constraint cumulativity

Without being confronted with experimental data, the idea of cumulativity might seem like an unnecessary complication of the theoretical model. However, once we have the richness of measurements of gradient acceptability at our disposal and look at them in charts such as Figure 2, Figure 3 and Figure 4, it becomes apparent that this notion makes the model more accurate. For example, if we look at the right half of Figure 3 in Catalan, it is plausible to think that whatever is responsible for the higher markedness of in situ focus with postverbal subject (38f) as opposed to its counterpart with preverbal subject (38e), is also responsible for the higher markedness of sentence-final focus with postverbal subject (38g) as opposed to its counterpart with preverbal subject (38h): the FULLINT constraint violated by postverbal subjects. Likewise, it is plausible to think that whatever is responsible for the higher markedness of the sentence-final c-focus position with preverbal subject (38g), as opposed to its counterpart with in situ c-focus (38e), is also responsible for the higher markedness of the sentence-final c-focus position with postverbal subject (38h) as opposed to its counterpart with in situ c-focus (38f): STAY-ϕ violated by *p*-movement. Then, we can elegantly explain the high degree of markedness of (38h) by assuming that the effects of violating FULLINT and STAY-ϕ simply add up.[20] Neither a standard-OT nor a stochastic OT approach can capture ef-

20. Furthermore, some constraints can be seen as interaction effects in the sense that their effect is due to the combination of other constraints. The idea is similar (though not identical, see below) to the concept of interaction in analysis of variance, where we do not only have main effects (e.g., *A* and *B*), but also an interaction ($A \times B$). Concretely, *[$_{\nu\text{P}}$ SUBJ ϕ FOC] and CP-VS can be described as interactions. The first one is the interaction of FullInt and Stay-ϕ, i.e., *[$_{\nu\text{P}}$ SUBJ ϕ FOC] is violated if both FULLINT, and STAY-ϕ are violated. The second one is the interaction of FULLINT, FOCP$_{\text{C-F}}$, and TOPICFIRST, i.e., CP-VS is violated if FULLINT, FOCP$_{\text{C-F}}$ (and eventually TOPICFIRST) are *not* violated. FULLINT is implicated in both interactions. It seems that subject position (FULLINT) has a genuine interaction effect with STAY-ϕ and FOCP$_{\text{C-F}}$, which also reflects the statistical interaction effect between subject position and focus position that was one of the first results of the analysis of variance documented above. However, there is a difference between the algebraic approach of linear equations and the statistical technique of analysis of variance concerning the notion of interaction: In the algebraic approach, the constraints *[$_{\nu\text{P}}$ SUBJ ϕ FOC] and CP-VS are treated like any other constraint. Furthermore, independence between them is not assumed, contrasting with analysis of variance, where all structural components shown in equation (40) in Section

fects of constraint cumulativity that come along with a concept of violation cost. They have not been conceived for modeling gradient acceptability data.

My aim is not to stop at a qualitative analysis of constraint cumulativity, but to try to go one step further and capture this idea in an algebraic model (see also Bergounioux et al. 2007 on mathematics and phonology). If constraints are assumed to have specific numerical violation costs, which also add up linearly in the case of violations of several constraints, and if we dispose of a continuous measure for the level of acceptability, then it is worth exploring whether the acceptability of a sentence can be derived by means of an *equation* based on the violations the sentence incurred.[21]

The work of Keller (2000: 256–265), who introduced the idea of calculating numerical constraint weights with systems of linear equations, greatly inspired this proposal. However, the present approach differs in several important points: (i) Keller (2000) deals with the problem of inconsistent systems of equations by approximation, applying the least squares estimation method. Furthermore, he embeds least squares estimation in an inferential-statistical and not heuristic approach (see below). (ii) In the present study, I build equations with which the absolute acceptability values are expressed directly.[22] Keller (2000) deals with acceptability *differences* between two sentences S_i and S_j: Equation 1 expresses $S_1 - S_2$, equation 2 expresses $S_1 - S_3$, etc. This is consistent with his method of measurement (magnitude estimation), where a value for a sentence is always measured proportionally to another stimulus. I work with graphic rating where absolute values are measured on a rating scale.

2.4 are truly independent. This means that an analysis of variance can result in an interaction $A \times B$ even in those cases in which we find neither main effect A nor main effect B. However, in the present algebraic approach, constraints represent aspects of grammatical structure that can be related to other aspects of grammatical structure (for example, if FULLINT, and STAY-ϕ are both violated, then *[$_{vP}$ SUBJ ϕ FOC] is also violated). Therefore, the constraints are not necessarily independent.

21. The present approach does not imply that grammar is somehow governed by numbers. A model does never reflect the real world – at least I adhere to an epistemology that takes this stance (see Armatte 2005 on the relation between model and reality). An anonymous reviewer pointed out that one reason why it would not be very plausible to assume a grammar governed by numbers, is the fact that quantifiable weights must also be representable in some pre-numerical ways, referring to cultures that do not primarily work with decimal numbers. The field of ethnomathematics has been devoted to the issue of mathematical modeling across cultures (see for example Eglash 1997; Ascher 2004; Chemillier 2007), and to the issue of the conception of ordinality and ranking (Guyer 2010).

22. (i) represents the core methodological difference and is the main point in this discussion. (ii) concerns a rather technical difference. It would be a misunderstanding to see it as a "deeper reason" for opting for the heuristic approach. One could also modify Keller's (2000) approach in order to work with absolute values.

4. Mathematical modeling of cumulative constraint violations

4.1. Heuristic methodology

Before starting to build a model, we need to make the basic methodological choice between an algebraic and a probabilistic/statistical approach.[23] In principle, both approaches represent feasible options, but they represent different methodologies in the sense that they try to give answers to different questions.

The statistical approach consists in looking at the variation of all single answers and trying to find a function that represents the best approximation to the set of data points. This is realized by algorithms known as maximal likelihood or least squares estimation (e.g., Strang 2009: 218–229). At the same time, the amount of variance explained (or the degree of the fit) is reported, usually by some correlation coefficient. The function – a non-linear or, the majority of studies, a linear regression – can also be used for predicting new data. The methodological framework is the same than the one in which analysis of variance and similar statistical tests are embedded: one works with a stochastic distribution, error probabilities and confidence intervals, as outlined in Section 2.4 and applied in Section 3.

The algebraic approach which I am adopting here, takes a different perspective. It builds on an exact function. If the system of equation is not solvable, one modifies the system itself to obtain a solvable one. Since it is not a statistical approach, one does not distinguish between an observed and a predicted value (technically, the model does not include so-called residuals which are the difference between an observed value and a value given by the model). One is not concerned with the problem of inference from a sample to the population, i.e., with the problem of statistical estimation. The perspective is a different one. The world is, so to speak, restricted to the sample or even more, to a single idealized case: One works with arithmetic means of the sample, ignoring the variance between the single judgments of the individuals.[24] This idealization is legitimate in the scope of the adopted methodology, to put it in Gigerenzer's (2008: 20) terms: "Heuristics are frugal – that is they ignore part of the information". Needless to say, the fact that the present algebraic approach builds on an exact function and not on approximation does not mean that it is somehow "more precise".[25]

23. I am most grateful to Uwe Schomburg, professor of technical mechanics at the Helmut-Schmidt-University in Hamburg, for long and inspiring discussions on the present mathematical approach. I owe important insights into modeling linear equations to him. All errors are of course mine.
24. One might also relate this approach to a description of E-grammar, contrasting with a usage-based model (see Newmeyer 2003 on this distinction).

Since the notion of a larger population is left aside, the mean values are not seen as estimations of the population, either. The question of generalization is not asked during model building (however, it is often, explicitly or implicitly, part of the ensuing discussion).[26] In this respect, the methodology is qualitative, and the method is quantitative.

It is a *heuristic* approach: The goal is not hypothesis testing but model building (or hypothesis generation). A methodology of doing mathematics for gaining insight and intuition, or discovering new patterns and relationships is often associated with the field of *experimental mathematics* (Borwein and Bailey 2004) where typically computational approaches are used extensively. This methodology is rooted in the work of theoretical mathematicians (e.g. Thom 1972, 1990), and of philosophers of science (Lakatos 1976, 1978). I am opting for the heuristic approach, because it can suggest new linguistic constraints within a framework of cumulative constraint violations.

Heuristic quantitative approaches are commonly used in multivariate data analysis, mainly in methods of group identification, and variable reduction such as cluster analysis (see Anderberg 1973), canonical correlation (Rao 1952), and factor analysis (see Harman 1976). These techniques are applied in psychology and sociology when one wants to uncover relevant constructs that underlie a quantitative pattern, which is too complex to be visible to the naked eye. Many of these methods can produce a set of theoretically infinite solutions.[27] The

---

25. Heuristic methodology takes a critical stance to the notion of precision, both regarding its usage and its standing in scientific research. Fields-laureate Thom (1983: 33) reproaches in his criticism of modern science a drive towards "fantastic numerical precision", as a result of what he calls "blind experimentation" coupled with a lack of emphasis on understanding. He goes on arguing that mathematical modeling can be a promising means to gain that deeper understanding which according to him, is not sufficiently in the focus of many scientists (see also Thom 1980). The notion of prediction is discussed in detail (note that one of his books, Thom 1993, is entitled "Predicting is not Explaining"). One core argument is that a function with a high degree of fit has not *per se* an explanatory value. It can in some cases even be counter-productive, because a function with a lower predictive power might trace the way to a better model.
26. Within this logic, the reason why one still works with a sample of a couple of dozen individuals and not with a single individual is to make a generalization *plausible*. The issue of generalization constitutes one important methodological debate, in particular in the context of approaches that do not build on inference (Bortz and Döring 1995: 310ff.). While some scholars argue against a prevalence of the notion of generalization putting emphasis on the notion of understanding (see also Footnote 25), others have proposed the idea of *exemplified generalization* (for example Wahl et al. 1982: 206ff.).
27. One should also bear in mind that an inherent characteristic of any quantitative modeling is that "one may refine indefinitely the model but the difference between the model and the real-world phenomenon, process or dynamical system which is subject to modeling cannot be avoided" (Pozna et al. 2010: 182). This argument can be traced back to the philosophy of science and scientific models (Bouleau 1999; Pullum and Scholz 2001; Armatte 2005). Heuristic modeling roots in a philosophy of science that is somewhat different from clas-

qualitative character resides in the fact that the researcher makes, after scrupulously weighting different quantitative and qualitative criteria, *a decision* in favor of a particular solution; in the present case it would be a decision in favor of a solvable system of linear equations, meaning a decision in favor of a certain constraint set.

One anonymous reviewer rightly stated that statistical approximation – as applied in Keller's (2000) work or any work that builds on linear regression analysis – *can* also be used in a heuristic approach for hypothesis generation (rather than hypothesis testing). The regression technique would then be used to evaluate models and suggest parameters. One could for example start with an initial model (i.e. an initial constraint set), solve the model by least squares estimation, evaluate its degree of fit or the error term, and eventually add/remove constraints in a heuristic way in view of a model with a higher fit. As a matter of fact, Zembowicz and Zytkow (1992) have done systematic work on this issue and proposed a discovery system called 'Equation Finder'. Its aim is to suggest what Ćwik, Koronacki and Żytkow (1999) call "acceptable models". The idea is to estimate $f(X)$ in the general regression (58), which is supposed to be the function underlying the data (see Moulet 1997 for a discussion of the Equation Finder). An interesting aspect of Zembowicz and Zytkow's (1992) work is that $f(X)$ is not restricted to a simple linear function but that it can incorporate a number of transformations and combinations of polynomials.

(58) $Y = f(X) + \varepsilon(X)$

In the present work, the choice in favor of the algebraic approach is not rooted in skepticism towards statistical analysis (which the present work also incorporates by applying analysis of variance). Rather, it allows posing the problem in a novel way, namely taking the inconsistency of the system of linear equations as a hint towards potentially missing constraints. The repertory of existing approaches to problems of modeling is large. Apart from least squares estimation and the proposed algebraic approach, other quantitative methods used in modeling include Support Vector Machines (Vapnik 1995), least squares Support Vector Machines (Suykens and Vandewalle 1999; Wen et al. 2008), Bayesian filtering (Pozna et al. 2010), to name just some of them. The point I want to make is that we are far from having a 'best method' in quantitative modeling – asking for the 'best method' might even not be a fruitful research question due to the very nature of heuristics.[28]

---

sic positivism (Lakatos 1976; see also Kiss 2006). This does not mean that evaluation and subsequent optimization are not important issues in heuristic mathematics.

28. Given the popularity of regression analysis among linguists, some words on this technique seem necessary (see Kruskal and Majors 1989; Budescu 1993; Bring 1994): First, one has to show great care in estimating the importance of the different predictors – which is accord-

To sum up, the reason why I choose an algebraic approach as outlined in the following section is not due to the fact that it is superior to other alternatives, but because it fits the way I approach the problem, namely that I idealize on random errors, excluding them by taking the mean value, and proposing what makes a system of linear equations unsolvable can be a missing constraint.

4.2. Basic notions of the approach

The core idea of the model of violation cumulativity is that each constraint comes with a specific numerical violation cost, and that the sum K_j of the violation costs of all constraints that a given sentence S_j violates corresponds to the amount by which the acceptability of S_j is lowered. Assume a simple case with three constraints c_1, c_2, and c_3. The violation of any of these constraints has a numerical violation cost of k_1, k_2, or k_3. Assume that the second candidate S_2 of a given set of sentences violates c_1 and c_3, but not c_2. Then

---

ing to Johnson (2000: 1) "one of the most common requests of statistical consultants and researchers". To put it in technical terms, one should not overestimate the stability of weight estimation (according to common nomenclature, b_i is the i-th weight parameter). Multiple regression builds on the idea shown in (i), according to which the total correlation R is the sum of *weighted* single correlations r_{ic} between each predictor (these could be constraints in the present case) and the dependent variable c (this could be the acceptability value of a sentence). However, due to what statisticians call *multicollinearity,* the estimation of the specific impact of a single predictor, its *importance* (Bring 1994), can hardly be isolated when predictors intercorrelate (which is mostly the case). (ii) illustrates this point. Each single correlation r_{ic} *contains* all other (weighted) single correlations. As a matter of fact, even small differences in the sample can cause large differences in regression weights (Johnson 2000: 3).

(i) $R = \sqrt{\sum_{i=1}^{k} b_i \times r_{ic}}$

(ii) $$\begin{array}{ccccccccccc} b_1 & + & b_2 r_{12} & + & b_3 r_{13} & + & \cdots & + & b_k r_{1k} & = & r_{1c} \\ \vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\ b_1 r_{k1} & + & b_2 r_{k2} & + & b_3 r_{k3} & + & \cdots & + & b_k & = & r_{kc} \end{array}$$

Second, simulation studies have shown that the same instability exists on a more general level, i.e., when one does not restrict oneself to linear regression but when non-linear relationships are also taken into consideration (Ćwik et al. 1999: 548). The latter means that one tries to find any function $f(X)$ by systematically searching among polynomials of transformed data such that the degree of fit is maximized for the relationship between predictors X and dependent variable Y (the equation is shown in (58)). As with linear regression, we observe instability, for example different random samples from the same distribution suggest substantially different functions $f(X)$.

K_2 corresponds to $k_1 + k_3$. We can also write $K_2 = 1 \times k_1 + 0 \times k_2 + 1 \times k_3$. The coefficients a_m of the violation costs k_m indicate whether the sentence violates the constraint ($a_m = 1$) or not ($a_m = 0$). As regards the example S_2, the coefficients for the three constraints are $a_{12} = 1$, $a_{22} = 0$, $a_{32} = 1$ (the first index refers to the respective constraint, the second index, here 2, refers to the sentence number).

Given a tableau with n constraints, the total violation cost K_j of a sentence S_j can therefore be written as a *linear equation in n variables* (or n unknowns), as in (59). This linear equation has two side conditions: (i) The coefficients can be only 0 or 1 because they reflect absence or presence of constraint violation, and (ii) violation costs cannot be negative, i.e., $k_m \in R^+$.

(59) $$K_j = \sum_{m=1}^{n} a_{mj} \times k_m, \text{ where } \alpha_{mj} \in \{0,1\} \text{ and } k_m \in R^+.$$

Violation costs must be positive, in order to reflect the decrease in acceptability for which the idea of constraint violation stands. Therefore, k_m represents the amount by which the line on the graphic rating scale is 'shortened'. In other words, violation costs are linearly represented on the scale (see Borg and Schönemann 1996 on the notion of scaling). A negative cost would mean that the violation of a constraint comes along with a "bonus" (as a matter of fact, Legendre et al.'s 1990 Harmonic Grammar do implement this idea; however, for a criticism see Prince 2003; Boersma and Pater 2008: Ch. 3.5; Pater 2009: Ch. 2.1). Note that in its present form (59) does not implement the idea of multiple violations of the same constraint (called "counting cumulativity" by Jäger and Rosenbach 2006). However, such an extension can be expressed in a straightforward manner by defining a_{mj} as a member of the set of non-negative integers, i.e., $a_{mj} \in N_0$ where a_{mj} expresses the number of times constraint c_m is violated in sentence S_j.

In order to indicate within the model of cumulative violation costs the acceptability value $\overline{X}_j$ (which is the mean value of the sample) of a sentence S_j, we need to assume an optimal value O. This optimum is supposed to be the acceptability value of a candidate in the hypothetical case that no constraint was violated. The acceptability value $\overline{X}_j$ corresponds to this optimum O minus the total violation cost K_j that the sentence S_j incurs. The acceptability of each construction corresponds to a linear equation as in (60).

(60) $$\overline{X}_j = 0 - K_j$$

Since each sentence corresponds to an equation as in (60), the entire tableau of the candidate set of q sentences and n constraints can be written as a *system of q linear equations with n unknowns* as in (61).

(61)
$$\begin{aligned}
\overline{X}_1 &= O - K_1 \\
\overline{X}_2 &= O - K_2 \\
\vdots & \qquad \vdots \\
\overline{X}_j &= O - K_j \\
\vdots & \qquad \vdots \\
\overline{X}_q &= O - K_q
\end{aligned}$$

The optimum O is constant for the entire candidate set, but the violation costs $K_1 \ldots K_j$ are variable. More precisely, the coefficients a_{mj} in equation (59) differ from sentence to sentence, which means that each sentence has a different pattern of constraint violations (all coefficients a form a $q \times n$ matrix A, where q is the number of sentences or linear equations, and n the number of constraints or unknowns).

This being said, two points need to be addressed: What does the optimum O correspond to? And under what conditions can such a system of linear equations be solved, in other words which limitations does it impose on us?

With regard to the optimum O, one could set it to the length of the graphic rating scale which is a fixed constant given by the test design. In the present study, O would then be 100. Setting O to the maximum of the scale would be appropriate if we worked with physical measurements. However, it is not appropriate when we work with human judgments. We know that arithmetic mean values of perceptive judgments fall into a smaller range somewhere within the physical rating scale, i.e., the theoretically possible range is not used. This is due to the psychometric phenomenon of the central tendency of judgments (see Section 2.3), as well as to the skewness of a unipolar distribution on the extremes (some subjects may assign the maximal value, some lower values, but no one can assign higher values). From a psychometric point of view and given that the following calculations build on mean values, it seems plausible to link the optimum O to the result of the item that scored highest rather than to the maximum of the physical scale.[29] It is of course necessary to ensure that the candidate set has been carefully designed such that potentially better candidates have not been forgotten. The optimum O is not a physical measure known before the study but, being an estimate of a perceptive variable, a value set post hoc.

29. Note that this choice also depends on the algebraic approach that builds on mean values. One might prefer linking the optimum to the scale maximum within an approach that builds on stochastic distributions such as least-squares estimation (if *single* measurements and not arithmetic means are the basic unit of analysis, the really operative scale range as mentioned in Section 2.3, spans the entire scale, since it is likely that a few single measurements hit the ceiling). I have to leave a more thorough discussion concerning the choice for the optimum O for future research.

One could set O equal to the score of the candidate with the highest acceptability value, i.e., $O = \max(\overline{X}_j)$. However, we need to take into account that our theoretical framework builds on the premise that all sentences violate some constraint, a premise that this approach shares with standard OT. Only the total violation costs differ. Since violation costs must be positive ($k_m \in R^+$), O must be greater than the highest acceptability value, i.e., $O > \max(\overline{X}_j)$. We can assume that the best candidate, i.e., the fully unmarked case, violates constraint(s) that have a negligible, or very low violation cost. Thus, I set this cost to a very low value, namely the smallest measurement unit in the graphic rating test, which is one percent of the scale length l.[30] In sum, the optimum of the candidate set O takes into consideration the psychometric nature of the test, the focus on mean values, and the theoretical premise of a negligible violation cost for the unmarked case. The fact that the optimum is linked to $\max(\overline{X}_j)$ also reflects the fact that the optimality of a sentence cannot be evaluated in isolation but only within a candidate set. It is calculated as follows.[31]

(62) $O = \max(\overline{X}_j) - u$, where u is the smallest measurement unit in the graphic rating test.

(62) expresses another side condition of the system of linear equations. If we substitute O in (60) by the right side of equation (62), bring K_m to the left hand side, and eliminate $\overline{X}_m$ we obtain (63); note that $\overline{X}_m = \max(\overline{X}_j)$, since (62) applies only to the equation with the highest acceptability value.

(63) $K_m = u$, where K_m is the total violation cost of the sentence S_m with the acceptability value $\overline{X}_m$, with $\overline{X}_m = \max(\overline{X}_j)$.

Given the present test design where $u = 1$, the total violation cost K_m equals to 1 for the constructions with the highest acceptability value in each candidate set. These are the constructions with in situ c-focus, (38e) for Catalan and (39e) for Spanish.

30. In methodological terms, this is a conservative approach meaning that the conditions under which an outcome is not interpreted, are strict. A different way to deal with this issue would consist in tying the threshold of negligible cost to data dispersion (for example, one could tie it to the standard deviation values given in Footnote 15).
31. (62) and (63) could be stated differently if another test design, in particular another measurement instrument, was applied. For example, in case of a rating scale with a substantially higher number of measurement units (compared to the present one which runs from 0 to 100), one could consider setting the negligible violation cost to a higher value than the smallest measurement unit. Or in case of a reversed scale with high values corresponding to low acceptability, one would set O to $\min(\overline{X}_j) + u$.

4.3. Challenges

Before addressing the difficulties in solving a system of linear equations with many equations and unknowns, I start with a simple system of equations as in (64) (we have seen many of those in school). We can solve it, by using the laws of elementary algebra in order to eliminate variables: First by solving the top equation for x in terms of y, then by substituting this expression for x in the bottom equation which gives $y = 1$, finally by substituting it back into the top equation for x which gives $x = 1.5$.

(64) $$\begin{aligned} 2x + 3y &= 6 \\ 4x + 9y &= 15 \end{aligned}$$

This system of equations avoids a number of challenges, because it has a unique solution, and we only need paper and pencil to solve it. However, the problems are often more complex in engineering, physics, chemistry, computer science, economics, and other disciplines in which linear algebra plays a prominent role; the same applies to the present linguistic problem of a model based on numerical violation costs.

The foremost challenge is to try to obtain a solvable system of linear equations. When, say, a mechanical engineer needs to build a system of linear equations for an empirical problem, she/he needs to perform prior theoretical work in order to achieve a meaningful and solvable system, e.g. some variables might be excluded from the model or new ones added, or some data excluded or new data included in order have less or more equations.

Challenge 1: a solution set with a single unique solution. If we seek concrete numbers, in this case exact violation costs for each constraint, we need a system with as many equations as unknowns. A system with fewer equations than unknowns has to be avoided which usually has infinitely many solutions (such solution sets can be useful, but not for our purpose). Likewise a system with more equations than unknowns has to be avoided which usually has no solution (unless an approach of statistical approximation is adopted). This requirement imposes on us the limitation of working only with tableaus with as many sentences as constraints.

Challenge 2: a consistent system of linear equations. However, same number of equations and unknowns does not automatically mean that there is a solution. In the case of empirical measurements, the major challenge is the *inconsistency* problem, which means that the equations do not possess a common solution. In these cases, the equations lead to a contradiction, e.g. to a statement such as 0 = 2. In order to achieve consistency, one needs a system of equation where the left-hand sides are linearly independent. Given two matrices A and B, the question is whether a unique matrix X exists such that $AX = B$ or $XA = B$ (for algeabraic details, see Lay 2005). A first theoretical model (based, for ex-

AUTHOR'S COPY | AUTORENEXEMPLAR

ample, on a presumed set of constraint violations in a tableau, or on some physical rules in an engineering problem) might seem plausible, but the calculation could show that it is inconsistent. Then, the model must be *tweaked* to make it solvable which is often an important step in the whole process (again, this applies to an algebraic and not to a statistical approach which would work with approximations). Still, one might not be able to find a consistent system of linear equations, at least one that is also conceptually satisfying. An important point is that not every empirical problem, not every tableau of sentences and constraints can be handled with a system of linear equations, and we will not know it before we have not tried to model it.

Challenge 3: side conditions. One might come up with a consistent system of linear equations which has a unique solution, but which does not respect possible side conditions. An important side condition in the case of constraint violation tableaus is $K_m \in R^+$ according to which violation costs must not be negative.

Once the conceptual work is done, how do we actually solve a system of linear equations? Systems of three or four equations can be readily solved by hand, but larger systems normally require computers that apply specific algorithms (usually a modified version of Gaussian elimination for matrix decomposition or an iterative method). The solutions to the two systems of linear equations in the present study were computed with the software MATLAB 7.4.

4.4. The heuristic character

Given the above-stated challenges, building a system of linear equations for empirical measurements does not work "out-of-the-box". As already stated, the major challenge is the inconsistency problem. We need to add another variable (henceforth called z) to the system that is either considered as a noise variable or interpreted like the other predefined variables, i.e. interpreted as a constraint. This approach requires a system that has one more equation than unknowns, because z will take the slot of the last unknown. The idea is to try to obtain a solvable system by adding the unknown z to one or more of the equations.[32] The heuristic character resides in z, in the fact that it is inserted (and interpreted) post hoc according to both a quantitative criterion included in an algorithm, and a criterion of interpretability. The unknown z can be seen as a "missing piece of a puzzle", and like in a real puzzle we only get an idea

32. An inferential-statistical approach would work very differently. Building on the distinction between observed and predicted values, one would essentially propose the closest possible *approximation* to the observed values that is consistent. Each acceptability measurement would be seen as a stochastic distribution.

of its shape once the whole rest is in place. On a technical level, there is some resemblance between the unknown z and what Potts et al. (2010) call margin of separation a in their linear programming model for harmonic grammar: They are arbitrary elements that help to make equations (or inequations in the case of Potts et al. 2010) solvable. However, the unknown z differs by its heuristic idea, i.e. by suggesting the hypothesis of a missing constraint.

How is the unknown z applied in solving systems for numerical violation costs? z is required to be as small as possible, which I also derive from the idea that it is the last missing piece. Concretely, the Euclidian distance $|z|$ is minimized, and z is added to the minimal number of equations (ideally z is only added to a single equation). The unknown z is introduced according to the following algorithm. Step (iv) is optional, i.e., it depends on the researcher whether there is the possibility of a second, more permissive pass, or not.

(65) Seek a consistent system of linear equations respecting all side conditions by:

- (i) Starting with $n = 1$
- (ii) Adding the smallest possible $|z|$ to n equations.
- (iii) Recursing over $n = n + 1$ going back to step (ii), if (a) the system of linear equations has not been solved or (b) z is smaller than the negated threshold of interpretability $z < -3u$
- (iv) Making once a new trial starting over at (i) but dropping the condition (iiib), if the system of linear equations has not been solved.

The unknown z can be interpreted in two different ways: z can be regarded as an irrelevant element of the quantitative model. We would then assume that we do not lose any relevant information if we simply ignore it; it is the noise and not the meaningful signal. Alternatively, one can also consider z to be a meaningful element, which is yet unexplained. I suggest including both alternatives with a threshold approach. If $|z|$ is below a threshold, it is not interpreted. If it is above this value, it is interpreted. The idea of a threshold of interpretability is often used in heuristic quantitative methods.[33] I set this threshold to $3u$, i.e., to three measurement units of the graphic rating test, which would correspond to a very small violation cost. Given that $u = 1$, z is considered as noise if $|z| < 3$, otherwise I seek an interpretation for it.

33. In factor analysis, for example, one tries to reduce a high number of items, e.g., the answers to 200 questions on a Likert scale, to a limited number of factors, e.g., 10 super-items or synthetic variables. One essentially calculates intercorrelation structures among the items. In order to interpret a factor, only those items whose correlation with the factor exceeds a certain threshold are taken into consideration. Depending on the nature of the research object and on the researcher, this threshold is usually somewhere between 0.3 and 0.7. If the correlation falls below the threshold, the item is not given sufficient importance for the interpretation of the factor.

Whenever one has to interpret post hoc a variable in a heuristic method, it is possible that there is no plausible interpretation. There are two ways to deal with such a situation. Either, one considers z in this case as noise, too (though the level of noise is higher than the threshold). Or one rejects the solution for z, going back to the algorithm in (65) further recursing over $n = n + 1$. Importantly, if z is interpreted as a constraint, the side conditions for constraints apply, i.e., $\{k_m, z\} \in R^+$. If z is seen as noise, it can also have a negative value. Note that a z which is smaller than the negated threshold of interpretability ($z < -3u$) cannot be a constraint (because it is negative), although the Euclidean distance is above the threshold of interpretability ($|z| \geq 3$). Condition (iiib) excludes this case. Only if no solution can be found otherwise, this condition can be dropped in the optional step (iv), and the algorithm would restart, this time accepting larger noise.

It commonly happens in such algebraic approaches that the system of equations cannot be solved. In this case, the only possibility is to come up with a modified set of constraints, i.e., modify the tableau. Of course, it is not always possible to change constraint definitions, which also build on previous work, and one might have to conclude that a specific tableau cannot be modeled with a deterministic system of linear equations.[34] These are the limitations of the approach. With regard to the tableau for Catalan (Section 4.5) it worked seamlessly, with regard to the tableau for Spanish (Section 4.6) it required a rather minor modification of the original constraint set.

What does the strategy of adding an unknown mean for the general methodology of this approach? Building systems of linear equations is then not only a method for determining the violation costs of a set of predetermined constraints, but it also becomes a way of *discovering* a new, missing constraint by assigning a meaning to z. It has already been mentioned that heuristic approaches can assist to identify constructs that underlie a quantitative pattern that is too complex to be grasped by the naked eye. Nuanced acceptability judgments of a set of sentences are of this type. We are able to carefully inspect judgment patterns such as the one in Figure 2 and formulate hypotheses on a set of constraints. But we still miss the level of detail and complexity inherent to the metrical nature of the judgments and the number of sentences involved. This is exactly where heuristic quantitative approaches come into play. This approach allows us to uncover elements that we might not have captured with the initial constraint set. It gives us the violation cost of an unknown constraint represented by z, which we then try to interpret (if z is not considered as noise).

34. Similar situations can also occur with other heuristic methods, e.g., one might not be able to find a solution to a factor analysis (for a variety of problems such as multiple loadings, lack of statistical reliability of factors, etc.). One can then try to modify the original item set and then reduce the modified set by a new factor analysis.

In the following I will first try to apply the deterministic approach as outlined above to the candidate set of Catalan sentences, and then to the candidate set of Spanish sentences.

4.5. Constraint violation costs for Catalan

Based on the idea that the sum of all constraint violation costs cumulate, we can model the markedness differences between the sentences (38a) to (38h) with seven constraints: CP-VS, CP-LEFTFOC, *[vPSUBJϕFOC], FULLINT, STAY-ϕ, *COPY$_{[+F]}$, and FOCP$_{\text{C-F}}$. The first column in Table 2 indicates the sentence number, and the second column the acceptability value obtained by the sentence (these values were already displayed in the line charts in Figure 3). The leftmost constraint in the table (CP-VS) is supposed to have the highest, the rightmost one the lowest violation cost. FocP$_{\text{c-F}}$ has the cost k_1, *Copy$_{[+F]}$ has the cost k_2, Stay-ϕ has the cost k_3, FullInt has the cost k_4, *[vPSubjϕFoc] has the cost k_5, CP-LeftFoc has the cost k_6, and CP-VS has the cost k_7. The sentences are ordered such that the one which violates the, presumably, lowest ranked constraint is in the first row and the one which violated the, presumably, highest ranked constraint is in the last row. This ordering represents the initial hypothesis in terms of ranking position, building on the analyses in Section 3. If the system of equations can be solved, the solution will not only indicate the real ranking but also the exact violation costs (thereby also the different distances between the constraints).

Table 2: Constraint table for Catalan

	$\overline{X}_j$	CP-VS	CP-LEFTFOC	STAY-ϕ	FULLINT	*COPY$_{[+F]}$	*[vPSUBJϕFOC]	FOCP$_{\text{C-F}}$
(38e)	69.2157							*
(38a)	65.8137					*		
(38f)	60.3235				*			*
(38g)	59.7353			*				*
(38b)	59.3235				*	*		
(38h)	46.7451			*	*		*	*
(38d)	45.2059		*		*	*		
(38c)	38.6176	*	*			*		

AUTHOR’S COPY | AUTORENEXEMPLAR

Based on (61), we obtain the following system of linear equations for the eight sentences and the seven constraints of the tableau.

(66) a. $69.2157 = O - k_1$ (equation to (38e), Table 2)
b. $65.8137 = O - k_2$ (equation to (38a), Table 2)
c. $60.3235 = O - k_4 - k_1$ (equation to (38f), Table 2)
d. $59.7353 = O - k_3 - k_1$ (equation to (38g), Table 2)
e. $59.3235 = O - k_4 - k_2$ (equation to (38b), Table 2)
f. $46.7451 = O - k_5 - k_4 - k_3 - k_1$ (equation to (38h), Table 2)
g. $45.2059 = O - k_6 - k_4 - k_2$ (equation to (38d), Table 2)
h. $38.6176 = O - k_7 - k_6 - k_2$ (equation to (38c), Table 2)

According to (62), $O = 70.2157$. According to (63), $k_1 = 1$. Applying the algorithm in (65), the above system can be solved by adding the unknown z to (66e).

(66e′) $59.3235 = O - k_4 - k_2 + z$ (equation to (38b), with z)

The solution to the system of linear equations, with (66e′) instead of (66e), unveils the violation costs k_m for each constraint c_m, and is given from highest to lowest in (67). We obtain $z = 2.402$. Given that $|z| < 3$, it is below the threshold of interpretability set above. It is therefore ignored in the model and treated as non-essential information or noise.

(67) CP-VS $\gg$ CP-LEFTFOC $\gg$ STAY-ϕ $\gg$ FULLINT
$k_7 = 15.5$ $k_6 = 11.7$ $k_3 = 9.5$ $k_4 = 8.9$

$\gg$ *COPY$_{[+F]}$ $\gg$ *[VPSUBJϕFOC] $\gg$ FOCP$_{\text{C-F}}$
$K_2 = 4.4$ $k_5 = 4.1$ $K_1 = 1$

We see that in Catalan CP-VS has a high violation cost, *COPY$_{[+F]}$, *[VPSUBJϕFOC] (and of course FOCP$_{\text{C-F}}$) have rather low violation costs, while FULLINT, STAY-ϕ, and CP-LEFTFOC are in an intermediate range.

4.6. Constraint violation costs for Spanish

A different version of CP-VS applies to Spanish than to Catalan according to the definition (53) in Section 3.2. It is not only violated by the topic-focus construction (39c). Unlike in the Catalan tableau, CP-VS is also violated by (39a) which only projects FocP.

The system of linear equations for the sentences (39a) to (39h) and the seven constraints reads as follows (according to (62): $O = 75.5943$; according to (63): $k_1 = 1$). Again, FocP$_{\text{c-F}}$ has the cost k_1, *Copy$_{[+F]}$ has the cost k_2, Stay-ϕ has the cost k_3, FullInt has the cost k_4, *[vPSubjϕFoc] has the cost k_5, CP-LeftFoc has the cost k_6, and CP-VS has the cost k_7.

Table 3: Initial constraint table for Spanish

	$\overline{X}_j$	CP-LEFTFOC	*[VPSUBJϕ FOC]	CP-VS	*COPY$_{[+F]}$	STAY-ϕ	FULLINT	FOCP$_{C\text{-}F}$
(39e)	74.5943							*
(39g)	69.7736					*		*
(39b)	61.8491				*		*	
(39f)	55.5755						*	*
(39a)	53.7925			*	*			
(39d)	52.7075	*			*		*	
(39h)	44.0000		*			*	*	*
(39c)	41.8302	*		*	*			

(68) a. $74.5943 = O - k_1$ (equation to (39e), Table 3)
b. $69.7736 = O - k_3 - k_1$ (equation to (39g), Table 3)
c. $61.8491 = O - k_2 - k_4$ (equation to (39b), Table 3)
d. $55.5755 = O - k_4 - k_1$ (equation to (39f), Table 3)
e. $53.7925 = O - k_7 - k_2$ (equation to (39a), Table 3)
f. $52.7075 = O - k_6 - k_2 - k_4$ (equation to (39d), Table 3)
g. $44.0000 = O - k_5 - k_3 - k_4 - k_1$ (equation to (39h), Table 3)
h. $41.8302 = O - k_6 - k_7 - k_2$ (equation to (39c), Table 3)

The system of linear equations (68a) to (68h) remains inconsistent. The algorithm in (65) does not succeed. It is therefore necessary to modify the tableau, in order to obtain a different system of linear equations. One of the constraints, namely CP-VS given in (53), comes in different versions: We have a "more permissive" version (proposed in Section 3.2 for Catalan), violated only in case the CP layer has more than one projection and V^0 dominates the subject, and a "less permissive" version, already violated with a CP layer of 1 projection and if V^0 dominates the subject (proposed in Section 3.2 for Spanish). I therefore suggest building a modified system of linear equations based on the more permissive version of CP-VS. It is in my opinion a tolerable modification, although it is obviously a trade-off. The markedness of (39a) cannot be explained by the permissive version of CP-VS, any more. However, we will see that the unknown z now comes into play in a more prominent manner. It can "compensate" to some degree the loss of conceptual clarity, by introducing a new theoretical element. Below is the modified tableau, which only differs from Table 3 in that (39a) does not violate CP-VS.

Table 4: Modified constraint table for Spanish

	$\overline{X}_j$	CP-LeftFoc	*[vPSubjϕ Foc]	CP-VS	*Copy$_{[+F]}$	Stay-ϕ	FullInt	FocP$_{C\text{-}F}$
(39e)	74.5943							*
(39g)	69.7736					*		*
(39b)	61.8491				*		*	
(39f)	55.5755						*	*
(39a)	53.7925				*			
(39d)	52.7075	*			*		*	
(39h)	44.0000		*			*	*	*
(39c)	41.8302	*		*	*			

Applying the algorithm in (65), the equations based on the modified tableau, where k_7 is not part of the equation (69e) concerning sentence (39a) any more, can now be solved: The algorithm is successful at its second recursion, and the unknown z is added to equations (69d) and (69e). The final system of equations is given below. Note that z has been added to equation (69e), i.e., we have reasons to believe that the markedness of (39a) is now explained differently. Although we were “forced” to use a different version of CP-VS, there might also be an advantage in the modification of the model: We now use exactly the same constraint set for Catalan than for Spanish, which makes the direct comparison of the results easier.

(69)
a. $74.5943 = O - k_1$ (equation to (39e), Table 4)
b. $69.7736 = O - k_3 - k_1$ (equation to (39g), Table 4)
c. $61.8491 = O - k_2 - k_4$ (equation to (39b), Table 4)
d. $55.5755 = O - k_4 - k_1 - z$ (equation to (39f), Table 4)
e. $53.7925 = O - k_2 - z$ (equation to (39a), Table 4)
f. $52.7075 = O - k_6 - k_2 - k_4$ (equation to (39d), Table 4)
g. $44.0000 = O - k_5 - k_3 - k_4 - k_1$ (equation to (39h), Table 4)
h. $41.8302 = O - k_6 - k_7 - k_2$ (equation to (39c), Table 4)

The solution to the system of linear equations unveils the following violation costs k_m for each constraint c_m, shown in (70). This time the value for z, added to the equations to (39a) and (39f), is above the threshold of interpretability: $z = 13.5$. Given that $|z| \geq 3$, we do not see it as noise but as a meaningful element. It is included among the constraints in (70), though still as a question mark.

(70) *[VPSUBJϕFOC] ≫ CP-VS ≫ ? ≫ CP-LEFTFOC
$k_5 = 20.3$ $k_7 = 16.3$ $z = 13.5$ $k_6 = 9.1$

≫ *COPY$_{[+F]}$ ≫ FULLINT ≫ STAY-ϕ ≫ FOCP$_{\text{C-F}}$
$k_2 = 8.3$ $k_4 = 5.5$ $k_3 = 4.8$ $k_1 = 1$

4.7. FOCUSPROMINENCE *and* FOCUSSALIENCY

We need to identify an unknown constraint suggested by the quantitative analysis, with a violation cost of at least medium size, ranked lower than CP-VS but higher than CP-LEFTFOC. It is violated by the sentences (39a) and (39f), because it was added to their respective equations. What property do (39a) and (39f) have in common? Or, what rule do these two sentences violate that the other sentences do not? Before getting to this point, I want to recall that one prominent characteristic of information packaging is the association of intonational prominence and focus. I again refer to Büring and Gutiérrez-Bravo's (2001) FOCUSPROMINENCE constraint in (35) (a reformulation of Zubizarreta's 1998: 21 FPR), which I repeat in (71).

(71) FOCUSPROMINENCE:
Focus is most prominent.

As it is formulated, (71) is inviolable. Its violation leads directly to ungrammaticality, if not ineffability or absolute ungrammaticality in the sense of Fanselow and Féry (2002: 268). Prosodic prominence essentially corresponds to the correct assignment of F-marking based on the principle that $[X]_F$ is prosodically more prominent than [Y].

I assume that prosodic prominence does not only have a categorical realization, based on the distinction ±F as reflected by (71), but also a gradient realization, based on the degree of intonational *saliency* of the phonetic realization of a focused constituent in its environment. If the saliency is insufficient, the construction is more or less marked, but not ungrammatical. I call this constraint FOCUSSALIENCY which has a violation cost of z = 13.5 and which takes the place of the question mark in (70). The idea is that a context in which intonational focus marking does not receive enough prominence leads to violation of (72).

(72) FOCUSSALIENCY:
Focus is intonationally salient.

(72) is in line with studies in segmental phonology on the role of perceptual saliency. To give one example, Côté (2007) proposed that the likelihood of schwa (ə) omission in French correlates with the relative salience of the sur-

rounding consonants, i.e. consonants become increasingly marked as they appear in contexts where the cues that permit listeners to detect their presence become weaker or fewer in the absence of schwa.[35]

What the Spanish constructions (39a) and (39f) have in common, is the fact that (i) the subject is adjacent to the c-focused object, and (ii) there is defocalized VP-material after the c-focus, namely the locative adjunct. The other constructions do not have both properties at the same time. With respect to (i), overt subjects are generally intonationally salient elements in Spanish, whether they are focus-prominent or not. It is not surprising that a number of word orders exist where nuclear stress must fall on the overt subject, i.e. where no other element can be focus (Zubizarreta 1999: 4233ff.). (39a) and (39f) are not among the constructions in which nuclear stress rigidly falls on the subject, but the fact that the subject is adjacent to the c-focused object nevertheless lowers the intonational saliency of the latter. As regards (ii), Catalan presumably allows more easily larger chunks of defocalized material and allows for freedom in the order of defocalized XPs, which Spanish does not as the examples (23) to (26b) further above have shown. Importantly, it is easier and more common in Catalan than in Spanish to fully deaccent defocalized material. Therefore the degree of intonational saliency of the c-focused object is lower in Spanish if other verbal constituents follow it because the latter are not deaccented. Neither of the criteria (i) and (ii) by itself seriously diminishes intonational saliency, but if the two co-occur, then its effect becomes manifest and FOCUSSALIENCY is violated. This also explains why FOCUSSALIENCY plays a role in the constraint set for Spanish but not for Catalan.

The difference between FOCUSPROMINENCE in (71) and FOCUSSALIENCY in (72) is the fact that the first is non-violable and being related to the focus feature F, syntactic in nature. The latter is violable, and being primarily related to the intonational properties of the sentence, phonological in nature. In terms of a model of cumulative constraint violation, (71) is ranked very high if not highest and thus associated with a very high violation cost. (72) is ranked lower and its violation is not so costly that it would make the sentence ungrammatical right away. The violation of FOCUSSALIENCY does not mean that FOCUSPROMINENCE is violated (as in the examples (39a) and (39f)), but the violation of FOCUSPROMINENCE necessarily comes along with the violation of FOCUSSALIENCY. The difference in violability between (71) and (72) can be related to work in which a distinction between two different kinds of constraints has been proposed (Müller 1999; Keller 2000; Coetzee 2004).[36]

35. This is why in morphological junctures the probability of schwa insertion increases, moving from /VC-CV/ (e.g., *la demande* [lad(ə)mãd] 'the request') to /VC-CCV/ or /VCC-CV/ (e.g., *une demande* [ynd(ə)mãd] 'a request').
36. In Müller's (1999) modification of standard OT, (71) would be a grammaticality constraint

4.8. Results in comparison

Figure 5 shows the values of the constraint violation costs for both Spanish and Catalan in a line chart. Note that the line for Catalan is interpolating from CP-VS to CP-LEFTFOC, because FOCUSSALIENCY is not included in the set of constraints.

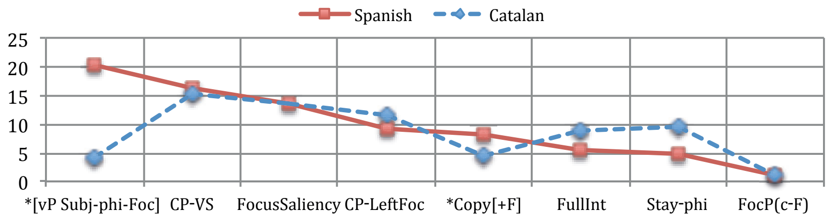

Figure 5: Constraint violation costs in comparison

In Spanish *COPY[+F] has a medium but STAY-ϕ a low violation cost. We find the opposite pattern in Catalan. Furthermore, *[VPSUBJϕFOC] shows the largest contrast. Its cost is very high in Spanish and low in Catalan. The hierarchy of the constraint violation costs for Spanish in (70) differs from Gabriel's (2007: 300) constraint ranking in (36). He identified ALIGNFOC, FULLINT, and STAY-ϕ as the main constraints that can account for word order variation. One should be prudent with a direct comparison because, unlike the constraints in (70), his hierarchy is not based on a mechanism of constraint cumulativity. However, it is interesting to note that FULLINT and STAY-ϕ reveal low in violation cost, and that ALIGNFOC is not included at all in the present analysis.

Let us turn again to the general debate on focus in Spanish (for which we find compared to Catalan more literature and opinions). It seems that the picture on focus and word order in Spanish might be more complex than what has been suggested. As regards i-focus, several authors (Hualde 2003: 160) accept $[S]_{\text{i-F}}$VO constructions or even consider it as the unmarked word order for narrowly i-focused subjects in Spanish. This contradicts the NSR for Romance (Zubizarreta and Vergnaud 2005) given in (10) above. Gabriel (2007) suggests that i-focus sometimes behaves like c-focus for which the NSR does not hold and for which the preposed position has been considered to be default. The differences in opinion on $[S]_{\text{i-F}}$VO constructions might be due to dialectal and

and (72) a markedness constraint. In Coetzee's (2004) rank-ordering model of EVAL (71) would be above the critical cut-off (that divides the constraint set into those constraints that a language is willing to violate resulting in markedness), while (72) is below it. In Keller's (2000) approach (71) would be a hard constraint and (72) a soft constraint.

microdialectal differences. However, another explanation is that at least some Spanish varieties allow different positions, possibly with nuances in markedness.

It is interesting that the systematic study of gradient acceptability judgments in the present work reveals that the preferred order for c-focus in the Spanish variety in Catalonia is *not* the preposed position. Rather, the in situ $SV[O]_{c\text{-}F}P$ and the sentence-final focus position $SVP[O]_{c\text{-}F}$ are the unmarked cases. One can consider the hypothesis that the Spanish variety of Catalonia works differently with respect to the position of c-focus. However, if we do not adopt this hypothesis, then we are observing a picture in which preposed i-focus and c-focus do behave differently but, somewhat ironically, in the opposite way than what has been suggested: The preverbal position can be the unmarked position for narrowly i-focused subjects and the marked position for c-focus. Future research can settle this issue by dialectal studies between Spanish varieties in which the type of focus (i-focus, c-focus) and the focused argument (subject, object) are systematically varied.

Another open issue is the motivation of the movement. While we have observed different degrees of markedness for in situ and preposed c-focus in Spanish (though no difference in terms of well- and ill-formed), the results on Catalan reveal that there is no acceptability contrast whatsoever between these two fully unmarked positions. Therefore, movement of c-focus in Catalan is well compatible with the idea of optional movement. I leave this question for future research, too.

5. Final remarks

I come back to the mathematical approach presented in the last section, discuss the potential contribution of such a heuristic methodology and lay out its limitations. Heuristic quantitative approaches are often applied when the question we are asking is centered on model building and on the most *plausible* solution. The issue of finding a constraint set within a framework of cumulative violation costs represents a good field of application for a heuristic approach. First, the idea of cumulativity makes the use of mathematical tools necessary. After a certain number of constraints (unknowns) and sentences (equations), it is beyond one's cognitive capacity to "see" how violations costs might add up to explain the degree of acceptability of a sentence. It is possible to conceive this problem as a mere computational challenge. Then the adequate answer consists in a statistical algorithm such as least squares estimation, which is a proven method for solving such cases. However, it is also possible to conceive this problem as a challenge of modeling and interpretation. In this case, the heuristic mathematical approach is the adequate answer in order to explore new concepts, possibly to create several solutions, which are compared and discussed.

The calculations in Sections 4.5 and 4.6 showed that this approach is not a magic tool. In some cases it works seamlessly like in the case of Catalan, in others it requires conceptual compromises like in the case of Spanish where we needed to apply a different version of the CP-VS constraint. And there might also be cases where it cannot be applied, because it would require some modification of the constraints that would lead to an implausible model. The main purposes of heuristic quantitative techniques is helping in finding satisfactory models and new insights, but it is not a device that just takes some input and produces some output. If one is aware of these limitations, this methodology is an enrichment of the empirical repertory of the linguist.

Finally, this work has shown the potential that lies in gradient data and concepts (in this case nuanced judgments and violation costs, but many other scenarios are conceivable in linguistics). Though we have witnessed in recent years a greater use of experimental methods for gathering fine-grained linguistic evidence, many theoretical concepts remain categorical. What seems promising to me is a direction in which we map as much as possible the richness of gradient data onto equally gradient concepts, and vice versa. Such a line of research comes with methodological challenges. A well-balanced combination of experimental techniques, inferential-statistical analyses, and heuristic approaches can prove useful here.

University of Freiburg
aria.adli@romanistik.uni-freiburg.de

References

Adli, Aria. 2005. Gradedness and consistency in grammaticality judgments. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives*, 7–25. Berlin: Mouton de Gruyter.

Adli, Aria. 2010. Constraint cumulativity and gradience: Wh-scrambling in Persian. *Lingua* 120(9). 2259–2294.

Albaum, Gerald, Catherine Roster, Julie H. Yu & Robert D. Rogers. 2007. Simple rating scale formats: Exploring extreme response. *International Journal of Market Research* 49(5). 633–650.

American Psychological Association. 2001. *Publication manual of the American Psychological Association*, (5th edn.). Washington: APA.

Anderberg, Michael R. 1973. *Cluster analysis for applications.* New York: Academic Press.

Antinucci, Francesco & Guglielmo Cinque. 1977. Sull'ordine delle parole in Italiano: l'emarginazione. *Studi di grammatica Italiana* 6. 121–146.

Armatte, Michel. 2005. La notion de modèle dans les sciences sociales: anciennes et nouvelles significations. *Mathématiques et sciences humaines* 172. 91–123.

Ascher, Marcia. 2004. *Mathematics elsewhere: An exploration of ideas across cultures.* Princeton: Princeton University Press.

Barbosa, Pilar. 2001. On inversion in Wh-questions in Romance. In Aafke C. J. Hulk & Jean-Yves Pollock (eds.), *Subject inversion in Romance and the theory of universal grammar*, 20–90. Oxford & New York: Oxford University Press.

Bard, Ellen Gurman, Dan Robertson & Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(1). 32–68.

Berger, Adam, Vincent J. Della Pietra & Stephen Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational Linguistics* 22(1). 39–71.

Bergounioux, Gabriel, Maïtine Bergounioux, Noël Nguyen & Sophie Wauquier. 2007. Quel programme de recherche en mathématique et phonologie? *Mathématiques et sciences humaines* 180. 9–26.

Bernardini, P. & Susanne Schlyter. 2004. Growing syntactic structure and code-mixing in the weaker language: The Ivy-Hypothesis. *Bilingualism: Language and Cognition* 7(1). 49–69.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32(1). 45–86.

Boersma, Paul & Joe Pater. 2008. Convergence properties of a gradual learning algorithm for Harmonic Grammar. New Jersey: Rutgers Optimality Archive. http://roa.rutgers.edu/files/970-0508/970-BOERSMA-0-0.PDF.

Bonet, Eulàlia. 1990. Subjects in Catalan. *MIT Working Papers in Linguistics* 13. 1–26.

Bonet, Sebastia & Joan Solà. 1986. *Sintaxi Generativa Catalana.* Barcelona: Enciclopèdia Catalana.

Borg, Ingwer & Peter H. Schönemann. 1996. Skalierung. In Edgar Erdfelder, Rainer Mausfeld, Thorsten Meiser & Georg Rudinger (eds.), *Handbuch quantitative Methoden*, 467–476. Weinheim: Beltz, Psychologie Verlags Union.

Bortz, Jürgen & Nicola Döring. 1995. *Forschungsmethoden und Evaluation für Sozialwissenschaftler.* Berlin & Heidelberg: Springer.

Bortz, Jürgen & Christof Schuster. 2010. *Statistik für Human- und Sozialwissenschaftler*, 7. Auflage. Berlin & New York: Springer.

Borwein, Jonathan & David Bailey. 2004. *Mathematics by experiment: Plausible reasoning in the 21st century.* Wellesley: A.K. Peters.

Bouleau, Nicolas. 1999. *Philosophies des mathématiques et de la modélisation: Du chercheur à l'ingénieur.* Paris: L'Harmattan.

Boye, Kasper & Elisabeth Engberg-Pedersen (eds.). 2010. *Language usage and language structure.* Berlin & New York: Mouton de Gruyter.

Bring, Johan. 1994. How to standardize regression coefficients. *American Statistician* 48(3). 209–213.

Budescu, David V. 1993. Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin* 114(3). 542–551.

Büring, Daniel & Rodrigo Gutiérrez-Bravo. 2001. Focus-related word-order variation without the NSR. In James McCloskey (ed.), *Syntax and Semantics at Santa Cruz, Vol 3*, 41–58. Santa Cruz: University of California.

Campos, Héctor & Mary Zampini. 1990. Focalization strategies in Spanish. *Probus* 2. 47–64.

Carlsson, Anna Maria. 1983. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. *Pain* 16(1). 87–101.

Casielles-Suarez, Eugenia. 1997. *Topic, focus and bare nominals in Spanish*. PhD dissertation, GLSA.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.

Chemillier, Marc. 2007. Mathématiques de tradition orale. *Mathématiques et sciences humaines* 178. 11–40.

Chomsky, Noam. 1971. Deep structure, surface structure and semantic interpretation. In Danny D. Steinberg & Leon A. Jakobovits (eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology*, 183–216. Cambridge: Cambridge University Press.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper and Row.

Coetzee, Andries W. 2004. *What it means to be a loser: Non-optimal candidates in optimality theory*. PhD dissertation, University of Massachusetts.

Coetzee, Andries W. 2006. Variation as accessing ‚non-optimal' candidates. *Phonology* 23(3). 337–385.

Cohen, Jacob. 1973. Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement* 33. 107–112.

Cohen, Jacob. 1988. *Statistical power analysis for the behavioral sciences (2nd edition).* Hillsdale: Erlbaum.

Contreras, Helen. 1991. On the position of subjects. In Susan D. Rothstein (ed.), *Perspectives on phrase structure: Heads and licensing* (Syntax and Semantics 25), 63–79. New York: Academic Press.

Costa, João. 1998. *Word order variation*. PhD dissertation, University of Leiden.

Côté, Marie-Hélène. 2007. Empty elements in schwa, liaison and h-aspiré: the French holy trinitiy revisited. In Jutta M. Hartmann, Veronica Hegedus & Henk van Riemsdijk (eds.), *Sounds of silence: Empty elements in syntax and phonology*, 59–101. Amsterdam: Elsevier.

Crawford, L. Elizabeth, Janellen Huttenlocher & Peder Hans Engebretson. 2000. Category effects on estimates of stimuli: Perception or reconstruction? *Psychological Science* 11(4). 280–284.

Ćwik, J., J. Koronacki & J. M. Żytkow. 1999. On discovering functional relationships when observations are noisy: 2D Case. *Lecture Notes in Computer Science* 1609. 547–555.

Diesing, Molly. 1992. *Indefinites.* Cambridge & London: MIT Press.

Domínguez, Laura. 2004. *Mapping focus: The syntax and prosody of focus in Spanish*. PhD dissertation, Boston University.

Eglash, Ron. 1997. When math worlds collide: Intention and invention in ethnomathematics. *Science, Technology, & Human Values* 22(1). 79–97.

Eguren, Luis & Olga Fernández-Soriano. 2004. *Introducción a una sintaxis minimista.* Madrid: Gredos.

Erdfelder, Edgar. 2010. Experimental psychology: A note on statistical analysis. *Experimental Psychology* 57(1). 1–4.

Erdfelder, Edgar & Jürgen Bredenkamp. 1994. Hypothesenprüfung. In Theo Herrmann & Werner H. Tack (eds.), *Methodische Grundlagen der Psychologie* (Enzyklopädie der Psychologie, Themenbereich B Methodologie und Methoden, Serie I. Forschungsmethoden der Psychologie. Band 1), 604–648. Göttingen, Bern, Toronto & Seattle: Hogrefe.

Erteschik-Shir, Nomi. 1986. Wh-questions and focus. *Linguistics and Philosophy* 9(2). 117–149.

Estebas-Vilaplana, Eva. 2000. *The use and realisation of accentual focus in Central Catalan with a comparison to English*. PhD dissertation, University College London.

Estebas-Vilaplana, Eva. 2003. Tonal structure of post-focal L in English and Catalan. *Atlantis* 25(2). 39–53.

Face, Timothy L. 2002. Local intonational marking of Spanish contrastive focus. *Probus* 14(1). 71–93.

Fanselow, Gisbert & Caroline Féry. 2002. Ineffability in grammar. In Gisbert Fanselow & Caroline Féry (eds.), *Resolving conflicts in grammars: Optimality theory in syntax, morphology, and phonology* (= Linguistische Berichte, Sonderheft 11), 265–307. Hamburg: Buske.

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang & Axel Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, Instruments & Computers* 39(2). 175–191.

Featherston, Sam. 2009. A scale for measuring well-formedness: Why syntax needs boiling and freezing points? In Sam Featherston & Susanne Winkler (eds.), *The fruits of empirical linguistics*, Volume 1: *Process*, 47–74. Berlin: Walter de Gruyter.

Featherston, Sam & Susanne Winkler (eds.). 2009. *The fruits of empirical linguistics*, Volume 1: *Process*. Berlin: Walter de Gruyter.

Feldhausen, Ingo. 2008. *The prosody–syntax interface in Catalan*. PhD dissertation, Universität Potsdam.

Fernández Soriano, Olga. 1989. Strong pronouns in null-subject languages and the avoid pronoun principle. *MIT Working Papers in Linguistics* 11. 228–239.

Féry, Caroline, Gisbert Fanselow & Manfred Krifka (eds.). 2007. *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS) 6.* Potsdam: Universitätsverlag Potsdam.

Gabriel, Christoph. 2007. *Fokus im Spannungsfeld von Phonologie und Syntax: Eine Studie zum Spanischen.* Frankfurt: Vervuert.

Gigerenzer, Gerd. 2008. Why heuristics work. *Perspectives on Psychological Science* 3(1). 20–29.

Goodall, Grant. 1991. Spec of IP and Spec of CP in Spanish Wh-questions. In M. Mithun, G. Perissinotto & Eduardo P. Raposo (eds.), *Linguistic perspectives on the Romance languages*, 199–209. Amsterdam & Philadelphia: John Benjamins.

Goodall, Grant. 2001. The EPP in Spanish. In W. Davis & S. Dubinsky (eds.), *Objects and other subjects*, 193–223. Dordrecht: Kluwer.

Grimshaw, Jane. 1997. Projections, heads, and optimality. *Linguistic Inquiry* 28. 373–422.

Grimshaw, Jane & Vieri Samek-Lodovici. 1998. Optimal subjects and subject universals. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough?*, 193–219. Cambridge: MIT Press.

Grosjean, F. 1982. *Life with two languages.* Cambridge: Harvard University Press.

Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents.* Dordrecht: Foris Publications.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation.* Cambridge: Cambridge University Press.

Gutiérrez-Bravo, Rodrigo. 2002. Focus, word order variation and intonation in Spanish and English. In Caroline R. Wiltshire & Joaquim Camps (eds.), *Romance phonology and variation* (Selected papers from the 30th Linguistic Symposium on Romance Languages, Gainesville, Florida, February 2000), 39–53. Amsterdam: John Benjamins.

Gutiérrez-Bravo, Rodrigo. 2006. *Structural markedness and syntactic structure: A study of word order and the left periphery in Mexican Spanish.* New York & London: Routledge.

Gutiérrez-Bravo, Rodrigo. 2007. Prominence scales and unmarked word order in Spanish. *Natural Language & Linguistic Theory* 25. 235–271.

Gutiérrez-Bravo, Rodrigo. to appear. La identificación de los tópicos y los focos. *Nueva Revista de Filología Hispánica*.

Guyer, Jane I. 2010. The eruption of tradition?: On ordinality and calculation. *Anthropological Theory* 10(1-2). 123–131.

Halliday, Michael A.K. 1967. Notes on transititivity and theme in English II. *Journal of Linguistics* 3. 199–244.

Hamblin, Charles Leonard. 1958. Questions. *Australasian Journal of Philosophy* 36. 159–168.

Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10. 41–53.

Harman, Harry H. 1976. *Modern factor analysis*, 3rd Edition. Chicago: University of Chicago Press.

Hayes, Mary H. S. & Donald G. Patterson. 1921. Experimental development of the graphic rating method. *Psychological Bulletin* 18(2). 98–99.

Hollingworth, Harry Levi. 1910. The central tendency of judgment. *Journal of Philosophy, Psychology, and Scientific Methods* 7. 461–469.

Horvath, Julia. 2000. Interfaces vs. the computational system in the syntax of focus. In Hans Bennis, Martin Everaert & Eric Reuland (eds.), *Interface strategies*, 183–206. Amsterdam: Holland Academic Graphics.

Hualde, José Ignacio. 2003. El modelo métrico y autosegmental. In Pilar Prieto (ed.), *Teorías de la entonación*, 155–184. Barcelona: Ariel.

AUTHOR'S COPY | AUTORENEXEMPLAR

Hulk, Aafke & Natascha Müller. 2000. Crosslinguistic influence at the interference between syntax and pragmatics. *Bilingualism: Language and Cognition* 3. 227–244.

IDESCAT. 2007. *Anuario estadístico de Cataluña 2007.* Barcelona: IDESCAT (Institut d'Estadística de Catalunya).

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar.* Cambridge: MIT Press.

Jäger, Gerhard. 2007. Evolutionary Game Theory and Typology: A Case Study. *Language* 83(1). 74–109.

Jäger, Gerhard & Anette Rosenbach. 2006. The winner takes it all – almost: cumulativity in grammatical variation. *Linguistics* 44(5). 937–971.

Johnson, Jeff W. 2000. A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research* 35(1). 1–19.

Karimi, Simin. 2005. *A minimalist approach to scrambling: Evidence from Persian.* Berlin: Mouton de Gruyter.

Karttunen, Lauri. 1974. Presuppositions and linguistic context. *Theoretical Linguistics* 1. 181–194.

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

Kiss, Katalin É. 1998. Identificational focus versus information focus. *Language* 74(2). 245–273.

Kiss, Olga. 2006. Heuristic, methodology or logic of discovery? Lakatos on patterns of thinking. *Perspectives on Science* 14(3). 302–317.

Kondo-Brown, Kimi. 2005. Differences in language skills: Heritage language learner subgroups and foreign language learners. *The Modern Language Journal* 89(4). 563–581.

Krifka, Manfred. 1991. A compositional semantics for multiple focus constructions. In Adam Zachary Wyner & Steven Moore (eds.), *Proceedings of SALT I*, 127–158. Ithaca: Cornell University Press.

Krifka, Manfred. 2004. The semantics of questions and the focusation of answers. In Chungmin Lee, Matthew Gordon & Daniel Büring (eds.), *Topic and focus: A cross-linguistic perspective*, 139–151. Dordrecht: Kluwer Academic Publishers.

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Working Papers of the SFB 632, Interdisciplinary Studies on Information Structure (ISIS) 6*, 13–56. Potsdam: Universitätsverlag Potsdam.

Kruskal, William H. & Ruth Majors. 1989. Concepts of relative importance in recent scientific literature. *American Statistician* 43(1). 2–6.

Kupisch, Tanja. 2007. Determiners in bilingual German–Italian children: What they tell us about the relation between language influence and language dominance. *Bilingualism Language and Cognition* 10(1). 57–78.

Ladd, D. Robert. 1980. *The structure of intonational meaning.* Bloomington: Indiana University Press.

Lakatos, Imre. 1976. *Proofs and refutations: The logic of mathematical discovery.* Cambridge: Cambridge University Press.

Lakatos, Imre. 1978. *Mathematics, science and epistemology: Philosophical papers*, Volume 2. Cambridge: Cambridge University Press.

Lay, David C. 2005. *Linear algebra and its applications*, third Edition. Boston: Addison Wesley.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic Grammar: a formal multilevel connectionist theory of linguistic well-formedness: theoretical foundations. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, 388–395. Hillsdale, NJ: Lawrence Erlbaum.

Liberman, Mark. 1975. *The intonational system of English*. PhD dissertation, MIT.

Lindholm, K. J. & A. M. Padilla. 1978. Language mixing in bilingual children. *Journal of Child Language* 5. 327–335.

Lindman, Harold R. 1974. *Analysis of variance in complex experimental designs.* San Francisco: W. H. Freeman & Co.

López, Luís. 1999. VP-Ellipsis in Spanish and English and the features of Aux. *Probus* 11. 263–297.

López, Luís. 2009. *A derivational syntax for information structure.* Oxford: Oxford University Press.

Martín-Butragueño, Pedro. 2005. La construcción prosódica de la estructura focal en español. In Gabriele Knauer & Valeriano Bellosta von Colbe (eds.), *Variación sintáctica en español: Un reto para las teorías de la sintaxis*, 117–144. Tübingen: Niemeyer.

Meisel, Jürgen M. 1986. Word order and case marking in early child language. Evidence from the simultaneous acquisition of two first languages: French and German. *Linguistics* 24. 123–183.

Meyer, Paula. 1972. Some observations on constituent order in Spanish. In J. Casagrande & B. Saciuk (eds.), *Generative Studies in Romance Languages*, 184–195. Rowley: Newbury House Publishers.

Moulet, M. 1997. Comparison of three inductive numerical law discovery systems. In Gholamreza Nakhaeizadeh & Charles C. Taylor (eds.), *Machine learning and statistics*, 293–317. New York: Wiley.

Müller, Gereon. 1999. Optimality, markedness, and word order in German. *Linguistics* 37(5). 777–818.

Newmeyer, Frederick J. 2003. Grammar is grammar and usage is usage. *Language* 79(4). 682–707.

Ordóñez, Francisco. 2000. *The clausal structure of Spanish: A comparative study.* New York: Garland.

Paradis, J. & F. Genesee. 1996. Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition* 18. 1–15.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33(6). 999–1035.

Pierrehumbert, Janet Breckenridge. 1980. *The phonology and phonetics of English intonation*. PhD dissertation, MIT.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker. 2010. Harmonic Grammar with linear programming: from linear systems to linguistic typology. *Phonology* 27(1). 77–117.

Pozna, Claudiu, Radu-Emil Precup, József K. Tar, Igor Škrjanc & Stefan Preitl. 2010. New results in modelling derived from Bayesian filtering. *Knowledge-Based Systems* 23(2). 182–194.

Pradilla, Miquel Àngel. 2001. The Catalan-speaking Communities. In M. Teresa Turell (ed.), *Multilingualism in Spain: Sociolinguistic and psycholinguistic aspects of linguistic minority groups*. Bristol: Multilingual Matters.

Prieto, Pilar. 2002a. Entonació. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez-Saldanya (eds.), *Gramàtica del català contemporani*, 393–462. Barcelona: Empúries.

Prieto, Pilar. 2002b. *Entonació. Models, teoria, mètodes.* Barcelona: Editorial Ariel.

Prince, Alan. 2003. Anything goes. In Takeru Honma, Masao Okazaki, Toshiyuki Tabata & Shinichi Tanaka (eds.), *A new century of phonology and phonological theory: A Festschrift for Professor Shosuke Haraguchi on the occasion of his sixtieth birthday*, 66–90. Tokyo: Kaitakusha.

Prince, Alan & Paul Smolensky. 1993. Optimality Theory: Constraint interaction in generative grammar. *Technical Report 2*. New Brunswick, New Jersey: Center for Cognitive Science, Rutgers University.

Pullum, Geoffrey K. & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retoré (eds.), *Logical aspects of computational linguistics: 4th international conference* (Lecture Notes in Artificial Intelligence, 2099), 17–43. Berlin: Springer Verlag.

Rao, Calyampudi Radhakrishna. 1952. *Advanced statistical methods in biometric research.* New York: Wiley.

Reips, Ulf-Dietrich & Frederik Funke. 2008. interval-level measurement with visual analogue scales in internet-based research: VAS generator. *Behavior Research Methods* 40(3). 699–704.

Revert-Sanz, Vicente. 2001. *Entonación y variación geográfica en el español de América.* València: Universitat de València.

Rizzi, Luigi. 1982. *Issues in Italian syntax.* Dordrecht: Foris.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.

Rochemont, Michael S. 1986. *Focus in generative grammar.* Philadelphia: Benjamin.

Rooth, Mats. 1985. *Association with focus*. PhD dissertation, University of Massachusetts.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116.

Samek-Lodovici, Vieri. 2001. Crosslinguistic Typologies in Optimality Theory. In Géraldine Legendre, Jane Grimshaw & Sten Vikner (eds.), *Optimality-theoretic syntax*, 315–353. Cambridge, Massachusetts: MIT Press.

Schütze, Carson T. 1996. *The empirical base of linguistics: grammaticality judgments and linguistic methodology.* Chicago: University of Chicago Press.

Schwarzschild, Roger. 1999. GIVENness, AVOIDF and Other Constraints on the Placement of Accent. *Natural Language Semantics* 7(2). 141–177.

Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure.* Cambridge: MIT Press.

Selkirk, Elisabeth O. 1995. Sentence prosody: Intonation, stress, and phrasing. In John Goldsmith (ed.), *The handbook of phonological theory*, 550–569. Oxford: Blackwell.

Selkirk, Elisabeth O. 2002. *Contrastive FOCUS vs. presentational focus: Prosodic evidence from right node raising in English.* Paper presented at ‚Speech Prosody 2002: Proceedings of the 1st International Conference on Speech Prosody', Aix-en-Provence.

Sheehan, Michelle L. 2007. *The EPP and Null Subjects in Romance*. PhD dissertation, Newcastle University.

Solà, Jaume. 2002. Modificadors temporals i aspectuals. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez-Saldanya (eds.), *Gramàtica del català contemporani*, 2867–2936. Barcelona: Empúries.

Strang, Gilbert. 2009. *Introduction to linear algebra*, 4th edn. Wellesley: Wellesley-Cambridge Press.

Suñer, Margarita. 1994. V-movement and the licensing of argumental *wh*-phrases in Spanish. *Natural Language and Linguistic Theory* 12(2). 335–372.

Suykens, Johan A. K. & J. Vandewalle. 1999. Least squares support vector machine classifiers. *Neural Processing Letters* 9(3). 293–300.

Szabolcsi, Anna. 1981. The semantics of topic-focus articulation. In Jeroen Groenendijk, Theo Janssen & Martin Stokhof (eds.), *Formal methods in the study of language*, 513–541. Amsterdam: Mathematisch Centrum.

Szendröi, Kriszta. 2005. Focus movement (with special reference to Hungarian). In Martin Everaert & Henk van Riemsdijk (eds.), *The syntax companion*, Vol. 2, 404–431. Malden: Blackwell.

Thom, René. 1972. *Structural stability and morphogenesis.* Boston: Addison Wesley.

Thom, René. 1980. *Modèles mathématiques de la morphogenèse.* Paris: C. Bourgois.

Thom, René. 1983. *Paraboles et catastrophes.* Paris: Champs Flammarion.

Thom, René. 1990. *Semiophysics - a sketch.* Boston: Addison Wesley.

Thom, René. 1993. *Prédire n'est pas expliquer.* Paris: Champs Flammarion.

Truckenbrodt, Hubert. 2007. The syntax–phonology interface. In Paul De Lacy (ed.), *The cambridge handbook of phonology*, 435–456. Cambridge: Cambridge University Press.

di Tullio, Ángela. 1997. *Manual de gramática del español – Desarrollos teóricos: Ejercicios, Soluciones.* Buenos Aires: Edicial.

Turner, N. M., A. J. Van De Leemput, J. M. T. Draaisma, P. Oosterveld & O. T. J. Ten Cate. 2008. Validity of the visual analogue scale as an instrument to measure self-efficacy in resuscitation skills. *Medical Education* 42(5). 503–511.

Uszkoreit, Hans. 1987. *Word order and constituent structure in German.* Stanford & CA: CSLI Publications.

Vallduví, Enric. 1992. *The informational component.* New York & London: Garland.

Vallduví, Enric. 1993. Catalan as VOS: Evidence from information packaging. In William J. Ashby, Marianne Mithun, Giorgio Perissinotto & Eduardo Raposo (eds.), *Linguistic perspectives in the Romance languages*, 335–350. Amsterdam & Philadelphia: John Benjamins.

Vallduví, Enric. 1995. Structural properties of information packaging in Catalan. In Katalin É. Kiss (ed.), *Discourse configurational languages*, 122–152. Oxford: Oxford University Press.

Vapnik, Vladimir N. 1995. *The nature of statistical learning theory.* Berlin, Heidelberg & New York: Springer.

Wahl, Klaus, Michael-Sebastian Honig & Lerke Gravenhorst. 1982. *Wissenschaftlichkeit und Interessen. Zur Herstellung subjektivitätsorientierter Sozialforschung.* Frankfurt am Main: Suhrkamp.

Wen, Wen, Zhifeng Hao & Xiaowei Yang. 2008. A heuristic weight-setting strategy and iteratively updating algorithm for weighted least-squares support vector regression. *Neurocomputing* 71(16-18). 3096–3103.

Winkler, Susanne & Sam Featherston (eds.). 2009. *The fruits of empirical linguistics*, Volume 2: *Product*. Berlin: Walter de Gruyter.

Yip, Virginia & Stephen Matthews. 2000. Syntactic transfer in a Cantonese-English bilingual child. *Bilingualism Language and Cognition* 3(3). 193–208.

Zagona, Karen. 2002. *The syntax of Spanish.* Cambridge & Oxford: Cambridge University Press.

Zembowicz, Robert & Jan M. Zytkow. Year. Discovery of equations: Experimental evaluation of convergence. *Proceedings of the 10th National Conference on Articial Intelligence*, San Jose. AAAI Press. 70–75.

Zubizarreta, Maria Luisa. 1998. *Prosody, focus and word order.* Cambridge: MIT Press.

Zubizarreta, Maria Luisa. 1999. Tema y foco. In Iganacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, 4216–4244. Madrid: Espasa-Calpe.

Zubizarreta, Maria Luisa & Jean-Roger Vergnaud. 2005. Phrasal Stress, Focus, and Syntax. In Martin Everaert & Henk van Riemsdijk (eds.), *The syntax companion*, Vol. 3, 522–568: Blackwell.