FOCUSING ON FUNCTION WORDS: INTONATIONAL MEANS TO MARK VERUM FOCUS IN GERMAN AND FRENCH

1Giuseppina Turco, 2Christine Dimroth, 3Bettina Braun

1Max Planck Institute for Psycholinguistics, Nijmegen, NL

2University of Osnabrück, Germany

3University of Konstanz, Germany

Address for correspondence:
Giuseppina Turco
Max Planck Institute for Psycholinguistics,
Nijmegen, NL
Wundtlaan 1, 6525 XD, NIJMEGEN - The Netherlands
Tel.: +31 (0)24 3521 546 giusy.turco@mpi.nl
Abstract

Recent studies have shown that in French initial accents occur more frequently at the start of a focus domain. We investigated whether this tendency also holds when the focus domain starts with an auxiliary (since function words have a different prosodic status than content words). We therefore compared semi-spontaneous utterances in Verum Focus, which signals a contrast in response to a claim with opposite polarity (e.g., *The child IS eating the candies* in the context of a child *not* eating the candies) and predicate focus in French and German, two languages that differ in how prosody is used to mark information structure. Results showed that both groups differentiated the conditions prosodically, Germans more consistently than French speakers. In French, categorically distinct accent patterns were realized across contexts in half of the cases. Some of these accent patterns involved an emphatic initial accent on the auxiliary. The other half of the Verum focused verb phrases showed to be phonologically similar but phonetically different from non-Verum focus cases. Pragmatic differences are hence also signaled in French even though this is not so systematic as in German.

Keywords: Verum Focus, function word, German, French initial accent, dialogue-game
1. Introduction

The intonational realization of spoken utterances is highly variable. Nevertheless, linguists have succeeded in unraveling some of the factors that guide speakers’ choice for particular intonational realizations. One prime candidate that is known to influence speech prosody and intonation is information structure, which has been shown to change the position and types of pitch accents as well as prosodic phrasing (e.g., Bolinger, 1965, 1972; Chafe, 1974; Halliday, 1967; Jackendoff, 1972; Ladd, 1980).

Languages differ, however, in how basic pragmatic concepts such as topic and focus are linguistically encoded; linguistic means range from particles and syntactic operations to intonational realization (e.g., Büring, 1997, 2003; König, 1991; Ladd, 2008; Lambrecht, 1994; Vallduví, 1992). This paper takes a cross-linguistic perspective, comparing two typologically distinct, but regionally very close languages, German (Germanic language family) and French (Romance). They differ in a number of aspects such as phoneme inventory, rhythm, phonotactics, word order, to name just a few. Prosodically, French words are grouped into accentual phrases1, APs (Jun & Fougeron, 2000, 2002), which minimally consist of one content word plus one or more optional function words (Jun & Fougeron, 2000, 2002). The accentual realization of APs is phonologically rather restricted and the structural and pragmatic factors favoring one or the other realization are as yet poorly understood. One or more accentual phrases form an intonation phrase. In German, on the other hand, words are initially grouped into intermediate phrases (ip) (cf. Grice, Ladd, & Arvaniti, 2000) that form then an intonation phrase (IP). However, intermediate phrases in German may be comprised of longer stretches of speech than French APs and their accentual realization is less constrained than that of French APs.

---

1 This prosodic unit is known by many names, see Lacheret-Dujour and Beaugendre (1999) for a review.
In this paper we investigated how a strong pragmatic factor, contrast, affects the intonational realization of APs in French and ips in German. More specifically, we tested the prosodic status of function words to shed light on constraints and preferences of variability in French APs, compared to German. Undoubtedly, the encoding of information structure differs across languages and is dependent on the language’s phonological and morpho-syntactic structure (cf. Vallduví & Engdahl, 1996). While in the literature there is a lot of work on the linguistic marking of information structure (e.g., Baumann & Grice, 2006; Braun, 2006; Féry, 2001; Féry & Kügler, 2008; German & D’Imperio, 2010; Jun & Fougeron, 2000), most of these studies have investigated the effects of information structure on content words (primarily in referential expressions) (but see, e.g., Bosch, 1988; Kameyama, 1999; Karabanov, Bosch, & König, 2007; Venditti, Stone, Nanda, & Tepper, 2002 for effects of stressed and unstressed pronouns in German and English). The current experiments aim at a better understanding of the pragmatics-prosody interface by studying how information structure affects the accentual realization of function words like auxiliaries in contexts of Verum Focus (e.g., Féry, 2007; Höhle, 1992; Leonetti & Escandell-Vidal, 2009), i.e. a focus on the truth value of the proposition (e.g., The child IS eating the candies as a contrastive answer to a previous negation utterance The child is not eating the candies). The distinction between content words and function words does not only exist at a morpho-syntactic level of the language, but also at the phonological level (e.g., Berendson, 1986; Inkelas & Zec, 1993; Nespor & Vogel, 2007; Selkirk, 1984, 1986), which is why it plays a role in language acquisition (e.g., Christophe & Dupoux, 1996; Christophe, Guasti, Nespor, Dupoux, & van Ooyen, 1997; Gerken & McIntosh, 1993; Shady & Gerken, 1999) and speech processing (Braun, Dainora, & Ernestus, 2011). Moreover, the prosodic status of
function words differs across languages (e.g., Delais-Roussarie, 1999 for French; Monachesi, 1996 for Italian; Zec, 2005 for Serbian).

In what follows, we will give a brief overview of the prosodic representations of function words from a Prosodic Phonology perspective (cf. section 1.1), a general definition of Verum Focus (cf. section 1.2), an introduction to previous intonation studies on focus marking in German (cf. section 1.3) as well as French (cf. section 1.4). In section 1.5 we report the research questions and predictions on the basis on previous studies.

1.1 The prosodic status of function words

Studies framed in the tenets of the Prosodic Phonology theory (e.g., Nespor & Vogel, 2007; Selkirk, 1995b; Truckenbrodt, 1999, 2007) have examined the unique phonological behavior of function words with respect to that of content words. In many languages it is generally assumed that function words can surface either in a stressed, strong form (i.e. if focused or in phrase-final position) or in a stressless, weak form (i.e. in phrase-initial position), whereas content words can only appear in strong forms. The vast majority of function words with weak forms are monosyllabic; they are frequently found in everyday conversation. By contrast, content words do not have weak forms and this is why, unlike function words, they can form independent phonological words (or “prosodic words” - ω) (e.g., Booj, 1995; Peperkamp, 1996; Selkirk, 1995a). The assumption that only content words can be prosodic words has been argued to hold for many languages (Prince & Smolensky, 1993, among others), whereas the prosodic representation of function word is very complex and sensitive to the type of category function words belong to (e.g., Zec, 2005) as well as to cross-linguistic variation (e.g., Delais-Roussarie, 1999 for French; Monachesi, 1996 for
As far as German (e.g., Hall, 1999) and French (Delais-Roussarie, 1999) are concerned, non-focused function words followed by content words, i.e. \([je \ regarde]\) (“I look”), are analyzed as elements not belonging to the prosodic word of the content word (e.g. as “free” clitics, Selkirk, 1995a), i.e. \([(je) \ (regarde)_ω]\), and are integrated into the prosodic structure at a phrase level \((ϕ)\), i.e. \([(je) \ (regarde)_ω ϕ]\). This is shown by a series of arguments built on both segmental (i.e. for German, lack of syllabification between final-segment of the function word and initial-segment of the content word, Hall, 1999, and references therein) and suprasegmental properties (i.e. for French, absence of initial accent on a phrase-initial function word, Delais-Roussarie, 1999) of these combinations.

There are conditions, however, in which function words can be prosodic words on their own. One of these is focus (Selkirk, 1995a). In general, it is assumed that focus promotes prosodic word formation \((ω\)-formation), and this holds for function words too (e.g., Hall, 1999 for German; Selkirk, 1995b; Wennerström, 1993 for English; Zec, 2005 for Serbian). Focused function words in strong form (i.e., like the German strong form \([ιst]\) but not the weak form \([ιs]\), meaning “is”) can be independent prosodic words since they are stressed and therefore dominated by their own feet (which is then dominated by a prosodic word). As such, focused function words can be realized with a pitch accent (i.e. according to the Association of Pitch Accent Rule, Liberman, 1975; Pierrehumbert, 1980). This holds true for many languages like English (Selkirk, 1995a) and German (Hall, 1999). However, generative studies on the syntax-phonology interface have mainly built their theories and constraints (i.e. the alignment constraints, Selkirk, 1995b) on the basis of content words (but see, Truckenbrodt, 2007, for discussion on this issue). Then, it is not clear
up to what extent focused function words can have the same intonational properties (phrasing, etc.) that focused content words have and if so, whether there would be cross-linguistic differences. In other words, we know very little about the consequences of focus effects on the stressing and phrasing of functional elements.

For French, the prosodic status of function words has been addressed by taking into account mainly cases of non-focused function words which, in principle, are considered to be not stressable (*mots non accentogènes*) and which cannot form a phrasal unit on their own (i.e. *groupe intonatif*, Mertens, 1993). Yet, cases of focused monosyllabic and phrase-initial function words have not been analyzed in more formalistic work (Delais-Roussarie, 1999; Mertens, 1993) and not very systematically in more experimental work (e.g. Delais-Roussarie, 1995; Di Cristo & Chafcouloff, 1975; Mertens, 1993, see section 1.4). Therefore, it remains an open question whether in French focused phrase-initial monosyllabic function words can be promoted to receive a pitch accent and be phrased apart from the content word, or to be marked by some other intonational feature (i.e. initial accent) and still be part of the same accentual phrase of the content word they are syntactically attached to.

1.2 Verum Focus

Generally speaking, Verum Focus (e.g., Creswell, 2000; Höhle, 1992; Leonetti & Escandell-Vidal, 2009; Romero & Han, 2004) refers to cases of focus which can be rephrased as *it is true that* in response to a claim with the opposite polarity or to yes-no questions. Consider the assertion produced by a speaker who is describing a

---

2 Also called as *polarity focus* (e.g., Dik, 1981; Gussenhoven, 1983; Holmberg, 2007), as *auxiliary focus*, (cf. Hyman & Watters, 1984) or as *predicate/predication focus* (Güldemann, 2003).
3 The term Verum Focus was first used by Höhle (1992) for German cases involving emphasis on the finite verb. The author suggested that accenting the finite verb is one of the possible ways for conveying such a contrast. In other languages, this phenomenon can be realized with other linguistic means (i.e., with *do*-insertion in English, with *si* in French, etc.). Following Höhle, many authors (e.g., Engdahl, 2006; Portner, 2007; Romero & Han, 2004) have provided a detailed semantic analysis of Verum Focus and tried to cover in its definition the different cross-linguistics realizations of this phenomenon. In this paper we will adopt the broad notion of Verum Focus as a way to contrast the polarity of a whole sentence with its negated counterpart.
situation illustrated in a picture *In my picture the child is eating the candies.* If this utterance is produced with the intonation contour shown in (1), it not only asserts a fact (*a child eating the candies*), but also presupposes a contrastive alternative to the accented word (Braun & Tagliapietra, 2010; Jackendoff, 1972). In this example, one of the possible contrastive antecedents is a sentence with an opposite polarity, i.e. child *not* eating candies. Crucially in (1) this contrast is expressed by accenting a function word, the auxiliary.

(1)  

\[(In \text{ my picture}) \text{ the child IS eating the candies}\]

There are a number of ways how Verum focus can be realized. German is generally very flexible in its use of pitch accent types and their distribution. This allows speakers to accent every word in an utterance without making (additional) syntactic modifications (e.g., Vallduvi & Engdahl, 1996). As a consequence, words can be marked as contrastive without changing the order of the constituents (although certain word orders are better suited to express certain information structures, cf. Uszkoreit, 1986). In German, therefore, Verum Focus can be signaled quite easily via intonation, i.e., by realizing a pitch accent on the finite verb (Höhle, 1992). In addition, in Germanic languages, a special group of particles, like the German particles *doch/wohl* or the Dutch particle *wel*, roughly meaning “indeed” (Dik, 1981; Dimroth, Andorno, Benazzo, & Verhagen, 2010; Gussenhoven, 1983; Van Valin, 1975, among others), can be used in the same pragmatic context. Similar to the contrastive function fulfilled by the pitch accent on the finite verb, these particles have the role of signaling that a contrast has occurred with respect to an antecedent utterance. An example of such particles is shown in (2) below:
(2) *der hat sich dann entschieden, doch zu springen*

he has REFL then decided, indeed to jump

“He finally decided to jump nevertheless”

(as a contrastive utterance to an antecedent one like: "First he did not jump")

(Dimroth et al., 2010, p.10)

Direct equivalents of these particles used in Verum Focus contexts have not been attested in Romance languages (i.e. in French and in Italian, Dimroth et al., 2010). Rather, the same type focus is mostly realized via syntactic operations (i.e., called as Verum Focus Fronting, Leonetti & Escandell-Vidal, 2009) where the locus of the propositional emphasis is left-fronted and embedded under affirmative particles such as the French *si* (Krebat-Orecchioni, 2001) shown in (3) below:

(3) A: *Il ne fait pas beau*

“The weather is not nice”

B: *Si (il fait beau)*

“It IS nice”

(Krebat-Orecchioni, 2001, p.102)

One of the reasons why French uses predominantly (morpho-)syntactic operations to signal focus and contrast is because it is classified as *rightmost language*, which requires the main accent to occur on the rightmost content word of an accentual phrase (e.g. Ladd, 2008; Vallduvi, 1992; Vallduvi & Engdahl, 1996). This phonological (structural) constraint does not always allow French speakers to convey
certain pragmatic meanings unambiguously with intonation, a phenomenon also reported for other Romance languages (cf. Ladd, 2008; Rasier & Hiligsmann, 2007 for French; Swerts, 2007 for Romanian; Swerts, Krahmer, & Avesani, 2002 for Italian; Wagner & McCurdy, 2010 for French). Recent cross-linguistics experimental investigations (Rasier & Hiligsmann, 2007; Swerts, 2007; Swerts et al., 2002; Wagner & McCurdy, 2010) suggest that in Romance languages the relation between accent placement and information status of the words (new vs. given) within syntactic constituents is not so straightforward as it is in Germanic languages 4.

1.3 Prosodic marking of focus in German

In German, function words and content words alike can be accented for pragmatic or emphatic purposes, which may result in scope ambiguities when particles are concerned (Féry, 2006; Karagjosova, 2006; Krifka, 1998, 1999; Reis & Rosengren, 1997; Sudhoff, 2010; Sudhoff & Lenertová, 2006; Van Valin, 1975). When a word is focused (irrespective of whether it is a function or a content word), it always appears in a stressed strong form and can therefore be realized with a pitch accent (Hall, 1999). In broad focus contexts, German (like English and Dutch) places the nuclear accent on the internal argument of the intonation phrase (Birch & Clifton, 1995, 2002; Chomsky & Morris, 1968; Gussenhoven, 1984; Selkirk, 1984; Uhmann, 1991; Welby, 2003, hereafter: focus exponent). Such utterances (e.g., Der Mann hat ein BIER getrunken, “The man has drunk a BEER”) are often ambiguous between a narrow focus on the argument itself (answering the question What did he drink?) and a broad focus on the entire verb phrase (answering the question What did he do?) or even on the entire utterance (What happened?). According to Grice and Baumann (2005)

---

4 However, such a claim holds if we assume that focus is mainly signaled via accent placement. For discussions on this issue, see Ladd (2008, pp. 214-223). For a revaluation of the focal typology within Romance languages, see Face and D’Imperio (2005).
breadth of focus can be disambiguated by the accent type produced on the argument: $H^* L\%$ for non-contrastive contexts, $L^+H^* L\%$ for contrastive ones. However, it should be noted that certain information structure distinctions are signaled mainly by gradient, phonetic means (rather than by categorically different accent types). Therefore, the same phonological accent type may indicate both narrow and broad focus, but accents in narrow focus have been shown to be more emphatic, realized by higher and later peaks and by longer duration (Baumann, Grice, & Steindamm, 2006; Gussenhoven, 2002; Mehlhorn, 2001) and followed by unaccented post-focal elements (Baumann et al., 2006; Féry & Kügler, 2008). Contrastive and non-contrastive topics likewise are often distinguished by phonetic differences (peak scaling and alignment) and not necessarily by different accent types.

Another way to encode contrast in German is by realizing the so-called ‘hat pattern’ or ‘bridge accent’ (Büring, 1997; Mehlhorn, 2001; Wunderlich, 1991). The hat pattern in German has been associated with double contrast sentences with a contrastive topic and a contrastive focus (Mehlhorn 2001, Büring 1997). Phonologically, it is realized with a rising accent on the contrastive topic, a sustained high pitch, and a falling accent on the contrastive focus (Braun, 2006; Büring, 1997; Mehlhorn, 2001). However, there is little agreement on the phonological description of this contour in the literature (Baumann, Becker, Grice, & Mücke, 2007; Féry, 1993; Féry & Kügler, 2008; Kohler, 1991b; Krifka, 1998; Wunderlich, 1991), neither with respect to the prenuclear topic accent (either $H^*$ or $L^*+H$) nor with respect to the nuclear focus accent (either $H^*+L$, $!H^*+L$ or $L^*$). Moreover, the hat pattern is not always found in conditions with a double contrast (e.g., Braun 2006).
1.4 Prosodic marking of focus in French

In the French phonology literature, authors typically distinguish between the initial accent (Hi) marked by an F0 rise located near the left edge of the AP, and the final accent (H*), marked by an F0 rise consistently associated with the final full syllable of a content word at the right-edge of the AP. There is a considerable debate about the structural properties of these two accents. Some researchers (Di Cristo, 1998, 1999, 2000; Post, 2000) argue that both accents are pitch accents. Others (Jun & Fougeron, 2000, 2002; Welby, 2006) claim that the initial accent is a phrase accent whereas the final accent is a pitch accent. Compared to the final accent which has more stable acoustic properties such as longer rhyme duration (e.g., Fletcher, 1991; Pasdeloup, 1990; Rolland & Loevenbruck, 2002), a higher F0 value (Rolland & Loevenbruck, 2002), and a more consistent syllable-anchoring (Welby, 2006), the initial accent has less clear acoustic correlates (i.e. the temporal association of the initial rise is difficult to predict, Welby, 2006).

There are a number of factors that influence the occurrence and the location of initial accents in French (Jun & Fougeron, 2002; Pasdeloup, 1990, 1992; Rossi, 1985; Verluyten, 1984; Welby, 2006, among others). Some studies report that their distribution depends on the speaking rate (i.e., they are more likely to be realized in utterances with a slower speaking rate), the position of the word in the sentence (i.e. they are more frequent in sentence-initial position, Pasdeloup, 1990), the length of the AP (i.e. they are more frequent in APs made of 6 or more syllables, Jun & Fougeron, 2002; Welby, 2006), and speaker idiosyncrasies (Pasdeloup, 1990). Also, initial accents are more likely to be associated with content words than with function words (Delais-Roussarie, 1999; Garde, 1968; Hirst & Di Cristo, 1996; Jun & Fougeron,
2002; Séguiot, 1976; Welby, 2006). Under certain conditions, however, function words can receive an initial accent (Delais-Roussarie, 1995, 1999; Di Cristo & Chafcouloff, 1975; Mertens, 1993). For instance, when an AP contains a long sequence of unstressed clitics in a row, Hi can be realized on one of the function words, mainly for rhythmic reasons, i.e. to prevent a long distance between two AP-final accents without any intervening rhythmic element (e.g. Delais-Roussarie, 1995; Dell, 1984; Mertens, 1987; Pasdeloup, 1990). Accenting a function word can also depend on other factors like the category they belong to (i.e., negation adverbs, demonstrative, tonic and relative pronouns can get accented, Delais-Roussarie, 1995, 1999; Mertens, 1993, among others), their phrase-internal position (proclitics are never realized with an initial accent but enclitics can be pitch accented, e.g., Delais-Roussarie, 1999; Mertens, 1993) and the number of syllables (i.e., disyllabic function words are more likely to be accented than monosyllabic function words, Delais-Roussarie, 1995; Jun & Fougeron, 2002; Welby, 2006, among others).

Recent studies (Astésano, Bard, & Turk, 2007; German & D'Imperio, 2010) have found that pragmatic context represents a further predictor for the occurrence of initial rises in an AP. Astésano et al. (2007), for instance, investigated the prosodic realization of coordinating constructions like in 4(a)-(b):

4.a $[\text{les bagatelles}]_{N1} [\text{et les balivernes saugrenues}]_{N2+A}$

4.b $[\text{les bagatelles et les balivernes}]_{N1+N2} [\text{sauge dry}]_A$

(meaning: “Crazy trifles and nonsense”)

The authors noticed that the second NP (here: balivernes) attracted more initial rises.

---

5 The initial accent is also called as secondary stress, secondary accent (Fónagy, 1979; Pasdeloup, 1990), rhythmic regulatory stress (Delais-Roussarie, 1994), etc. The final accent is also referred to as primary stress, primary accent (Fónagy, 1979;
when the adjective had narrow scope (example a) than when it had broad scope (example b). On the other hand, the adjective attracted more initial rises when it had broad scope than when it had narrow scope. The initial rises hence appear to mark a strong syntactic boundary. Along the same lines, German and D’Imperio (2010) showed that the distribution of initial rises in the default pattern /LHiLH*/ is linked to both structural constraints (length of APs) and to information structural constraints (focus vs. non-focus). The authors claim that initial rises preferably occurred in APs whose left edge coincided with the left edge of a focus domain. Both of these studies are concerned with the prosodic marking of content words, however.

In contrastive emphatic contexts, French speakers usually realize a focal accent, $Hf$ (a single rising-falling contour) on the narrowly focused word (Clech-Darbon, Rebuschi, & Rialland, 1999; Di Cristo & Hirst, 1993; Hirst & Di Cristo, 1998; Rossi, 1985; Touati, 1987). There is experimental evidence (Jun & Fougeron, 2000) showing that such a focal accent can replace both the initial accent (Hi) as well as the final accent (H*) (Dohen & Loevenbruck, 2004; Jun & Fougeron, 2000), often depending on speaker’s choice (e.g., Dahan & Bernard, 1996; Jun & Fougeron, 2000), and that is usually accompanied by phrase breaks (e.g., Dahan & Bernard, 1996; Séguinot, 1976).

Post-focal elements are usually deaccented when they appear in a post-focal AP (Clech-Darbon et al., 1999; Hirst & Di Cristo, 1996; Jun & Fougeron, 2000; Touati, 1987); like other Romance languages, French speakers do not deaccent repeated (or given) words inside syntactic constituents (Ladd, 2008; Rasier & Hiligsmann, 2007; Swerts, 2007; Swerts et al., 2002; Wagner & McCurdy, 2010).6

---

6 An example of absence of anaphoric deaccentuation within syntactic constituents in French is the following: *Un flic AMÉRICAIN a rencontré un fermier AMÉRICAIN*, meaning: “An American policeman met an American farmer.” (Ladd, 2008; Wagner & McCurdy, 2010). However, Romance languages resist deaccentuation to different degrees, depending on the type of contrast (Ladd, 2008).
Rather, syntactic operations (i.e., right-dislocations) are considered to be the most common strategies for marking different focus structures. The pre-focal region tends to have a compressed pitch range (Dohen & Loevenbruck, 2004; Jun & Fougeron, 2000; Touati, 1987), a lengthened final syllable (Dohen & Loevenbruck, 2004) and is phrased as a single unit (e.g., Chen & Destruel, 2010; Delais-Roussarie, 1994; Dohen & Loevenbruck, 2004; Féry, 2001). Recent experimental work on the levels of prosodic constituency in French (D’Imperio & Michelas, 2010; Michelas & D’Imperio, 2010) has shown that an intermediate level of phrasing (ip) coincides with major syntactic boundaries (e.g., between a complex noun and the verb).

To sum up, initial accents are typically associated with one of the first syllables of the content word of the AP (Jun & Fougeron, 2002; Welby, 2006) and, although less clearly established, they can be associated with function words under certain conditions. However, clear pragmatic conditions affecting their prosodic status and experimental evidence attesting their relation with information structure are far from being well established. This study aims at filling this gap by comparing the intonational realization of Verum Focus in French and German.

1.5 Research questions and Hypotheses

In the present experiment, Verum Focus (VF) utterances like the B-sentences in 5(a)-(b) will be compared to similar utterances with Predicate Focus (which we will call non-VF henceforth), i.e. a context in which the focus encompasses the entire verb phrase (including the noun object) as in 6(a)-(b). In particular, we will investigate the prosodic means to disambiguate the pragmatic contexts by German and French speakers (the domain of focus is marked here by squared brackets with a subscripted $F$):
Our research questions are formulated as follows: What are the effects of focus on the phrasing and tonal realization of the auxiliary in German and French? In German, does the focused auxiliary receive a pitch accent compared to non-Verum Focus cases? In French, does the effect of focus trigger the presence of an AP-boundary to
the right of the focused function word (even if within the VP) or of the non-finite verb? Or, is the auxiliary realized with a left-aligned initial accent and is still part of the same AP of the non-finite verb?

Our predictions, derived from the literature reviewed above are as follows: In German, we expect a direct relationship between nuclear accent placement and focus exponent (the auxiliary for VF contexts and the object noun for non-VF contexts). More specifically, in non-VF cases, the internal argument (i.e. object noun) of the utterance should be associated with the nuclear falling accent (Birch & Clifton, 1995, 2002; Büring, 2006; Gussenhoven, 1984; Uhmann, 1991; Welby, 2003). In VF contexts, the nuclear accent is expected to be realized on the focused auxiliary. In addition to the targeted VF/predicate focus speakers might encode the contrast between the pictures they are talking about (In MY picture...). This might result in the production of hat patterns. However, no firm predictions are possible since previous experimental studies have either reported that the hat pattern is the prototypical contour for signaling double contrast (Mehlhorn, 2001), while Braun (2006) only found a few. Both analyses were based on read speech. We therefore investigate whether in the present semi-spontaneous productions, hat patterns are frequently used in double contrast conditions or not and which phonological form they have.

For French, there are two competing hypotheses. First, since the beginning of the focal domain coincides with the auxiliary in both contexts, the first prediction is that speakers produce an initial accent at the left edge of the focal domain in both conditions (German & D'Imperio, 2010). The initial accents could be either realized on a) the first syllables of the content word (as LHi, with the elbow L between the function and the content word), or b) as a plateau extending all over the syllable of the function word and be realized as an Hi without preceding L tone. For the same reason,
we do not expect differences in the pre-focal region of the sentence. Furthermore, the non-finite verb should not differ across contexts either because of lacking deaccentuation of anaphoric material within syntactic phrases in Romance languages (cf. section 1.2 and 1.4) (Ladd, 2008; Rasier & Hiligsmann, 2007; Swerts et al., 2002; Wagner & McCurdy, 2010). Consequently, the only intonational difference between contexts could be the accentual realization of the object noun: we expect accentuation of the object noun in non-VF but not in VF (Clech-Darbon et al., 1999; Di Cristo, 1998; Hirst & Di Cristo, 1996; Jun & Fougeron, 2000; Touati, 1987).

Second, given earlier studies on narrow focus and emphatic accent realizations in French (Clech-Darbon et al., 1999; Di Cristo, 1998; Rossi, 1985; Touati, 1987), another option is that function words (here the monosyllabic auxiliaries) are realized more frequently with an emphatic initial accent in VF than in non-VF contexts (see Jun & Fougeron, 2002, p.169 and footnote 2). This would imply that French speakers can signal the focal scope difference already within the verb phrase and not only in the post-verbal region (i.e. via post-focal deaccentuation of the object noun). If this was the case, the initial accent on the function word could either form a single AP on its own and be realized as a pitch accent or be phrased in an AP together with the non-finite verb and be realized as a phrase accent.

2. Methods

A picture-difference task between two speakers was designed for the controlled elicitation of Verum Focus utterances produced in a dialogue-game. For a more detailed description of this protocol see The Polarity-Switch Dialogue (Turco, 2009). This task represents picture comparisons in dialogue form between a confederate speaker and the participant and thus allows for a more spontaneous elicitation of
Verum Focus than possible in a reading task.

The structure of the dialogue-game is based on a 3-step schema: (a) a baseline picture, accessible to both speakers, in which a situation is illustrated (e.g. a child eating candies); (b) a negation picture where the opposite event is depicted (the child is not eating the candies), only accessible to the confederate; (c) an affirmation picture that is similar to the baseline picture (a) (the child is eating the candies), only accessible to the participant. The confederate speaker was instructed to keep eye contact and to produce very similar intonational realizations for the descriptions of his picture. By providing context negation sentences (Confederate: In my picture X is not doing Y), the task allowed us to elicit target sentences with the same word order but with a change on the polarity of the proposition (Verum Focus) (Participant: In my picture X is doing Y). All the target utterances contain a complex verb phrase (auxiliary + non-finite verb) so to address the question of whether speakers would locate the accent on the function word of the target phrase or not.

2.1 Participants

Eight French native speakers (2 male and 6 female, average age = 29.6 years, SD = 2.1) and eight German native speakers (2 male and 6 female, average age = 23.3, SD = 5.9) were recorded. The French speakers were master students at the University of Paris VIII and PhD students at the CNRS, others were researchers in Sociology at the CNRS in Paris. The German native speakers were all students at the University of Heidelberg. None of them had a reported history of speech/language impairment or other developmental deficits. They all received a small fee for their participation.
2.2 Materials

For each language group, a native speaker confederate produced the context utterances: a female speaker of standard Northern German (twenty-six years old) and a male French Parisian speaker (twenty-three years old). They were both trained to use the same intonation contour for all participants in a natural way (e.g., for VF condition: an accent on the German negation “NICHT” and on the French negation “ne…PAS”, for non-VF: in both languages an intonation contour with no particular emphasis on the whole context utterance was realized) across all the different trial conditions and for all participants.

PLEASE LISTEN TO CONFEDERATEFRENCH_EXAMPLE.WAV
PLEASE LISTEN TO CONFEDERATEGERMAN_EXAMPLE.WAV

There were 32 VF trials, 32 trials in a non-VF context, and 50 filler trials. The 32 VF trials consisted of 12 trials in which pictures illustrated present-perfect actions (auxiliary-condition; the target trials for the present investigation), 12 trials in which pictures illustrated simple-present actions (lexical-condition), and eight trials with emotional state pictures (copula-condition). The 32 non-VF trials had similar structures and the same type of verb phrase as the VF ones and served as control condition to test in which respect the observed intonation patterns are specific to the utterances elicited in the VF condition. As mentioned above (cf. section 1.5), both the VF utterances and the non-VF utterances contain a further initial contrast on the prepositional phrase (*In MY picture*...), resulting in a double contrast. This is due to the fact that participants had their own picture to compare to the other one. The experimental set-up did hence not elicit explicit corrections (Dik, 1981) but contrast,
because speakers talk about two different topic situations (i.e., in our case, represented by the two pictures, Klein, 2008).

For the research purposes of the present paper, we will focus our attention only on the auxiliary condition, whose materials are described in more detail here. For a description of the full set of materials see Turco (in prep). The auxiliary condition comprised 12 VF utterances and 12 non-VF utterances. All verbs were transitive and telic (e.g., Givón, 2001; Verkuyl, 1972) thus depicting completed actions. In all trials the auxiliary was expressed with the monosyllabic *a* in French and *hat* in German (“has”). In the corresponding pictures, the referent (the noun subject, in our case) was always represented by one repeated entity in all the three pictures of each trial. All stimuli are listed in the Appendix A. An example series of pictures is shown in Fig.1.

![Figure 1: Example of the Verum-Focus marking protocol.](image)

2.3 Procedure

Participants were instructed with a video-clip tutorial in the native language and familiarized with the task in a short warm-up session of 5 trials. The baseline pictures...
(cf. section 2) were displayed on an IBM screen, the individual pictures on two e-books, one for the confederate and one for the participant. Each picture described a single event and elicited only one target utterance. In the VF condition, the confederate had to describe the negation picture in comparison to the baseline picture, the participant had to describe the affirmation picture in comparison to the confederate’s description of the negation picture. The confederate had the first turn in all the mini-dialogue trials and, on comparing his picture with the baseline, could provide the negation context so as to cue the participant to the production of the VF structures. In the non-VF condition the procedure was the same, but the pictures showed two different actions performed by the same agent. Both interlocutors could not see each other's pictures and were not pressured by any time constraints. Moreover, in order to encourage participants to produce full sentences, they were told that the dialogue would later be used for another experiment, in which somebody else would have to listen to their descriptions presented in random order and would have to match them with corresponding pictures.

The French group was tested at the Department Structures Formelles du Langage UMR 7023 (CNRS) in Paris; the German group was tested at the University of Heidelberg. All sessions took place in experiment rooms and were sound-recorded using a Roland Edirol R-09 24bit digital wave/MP3 recorder and two phantom microphones (one for the confederate and one for the participant) SENNHEISER ME40 linked to a six channel audio mixer ALESIS 6FX. The microphones were placed at 30 cm distance from the speakers. For the contemporary display of the pictures on the three screens, computers were connected via (W)LAN connected to a 3COM-LAN Switch in a client/server configuration. The whole session, comprising of 110 trials, lasted approximately 20 minutes.
3 Results and Discussion

The 192 productions (96 for French and 96 for German) were first coded on the phrase, word and syllable level using Praat (Boersma & Weenink, 1992-2011). Trials in which participants produced very long pauses between phrases, utterances with hesitations or disfluencies were not included. Furthermore, we excluded sentences in which participants described the pictures with other words than the ones cued by the confederate (l’éléphant *a cassé le parquet* rather than *l’éléphant a défoncé le parquet*) or trials in which they described the actions with a tense other than the intended present perfect (*la fille est en train de déchirer le billet* rather than *la fille *a déchiré le billet*). In Table 1 we show the number of cases that have been analyzed out of a total of ninety-six items per condition and language:

<table>
<thead>
<tr>
<th>Types of items analyzed</th>
<th>Non-VF</th>
<th>VF</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>47</td>
<td>67</td>
<td>114</td>
</tr>
<tr>
<td>French</td>
<td>34</td>
<td>69</td>
<td>103</td>
</tr>
</tbody>
</table>

Table 1: Number of cases annotated for the two language groups in the two conditions.

Of the non-VF items, only 34 out of 96 cases could be analyzed for French and 47 for German: In this condition, participants were not expected to use the description produced already by the confederate but had to make up the sentences themselves based on the picture they saw. Many trials had to be excluded because participants used different tense and aspect, which would not have been comparable to the grammatical structure of the VF items.

The utterances were subsequently intonationally analyzed and labeled by adopting criteria of the Autosegmental-Metrical approach for both languages: GToBI
for German (Grice & Baumann, 2002; Grice et al., 2005) and the model described in Jun and Fougeron (2000, 2002) for French. The data were all labeled by the first author. In addition, forty percent of the dataset for each condition (both VF and non-VF) were also annotated by the third author to compute interrater reliability scores.

All the statistical analyses were performed by using the R software package (R Development Core Team, 2008).

3.1 German data

In German the interrater reliability score was 80.8% for the auxiliary hat (Kappa = 0.70 for the accentual realizations H*L-, H+L*, !H+L* and Unaccented) and 79% for the object noun (Kappa = 0.66 for the accentual realizations H*L-, H+L*, Unaccented).

3.1.1 Results

We will first describe the accent realizations of the auxiliary followed by the object noun and the non-finite verb for non-VF contexts and compare them to the intonation patterns produced in VF contexts.

In non-VF contexts, the auxiliary hat was generally not accented (98%) with the exception of one case, in which it was realized with a nuclear falling accent (H* L-). The object noun was mainly accented with a nuclear falling accent (H*L- or H+L*, four of which were downstepped: !H+L*); the non-finite verb was generally left unaccented. Only in one case the object noun was unaccented and followed by a pitch accented non-finite verb (H+L*).

7 There are many different descriptions on the intonation of German (Féry, 1993; Fox, 1984; Grice & Baumann, 2002; Grice et al., 2005; Kohler, 1991a; Mayer, 1995; Uhmann, 1991) and many other models proposed for the description of the French phonology (Di Cristo, 1998, 1999; Hirst & Di Cristo, 1996; Jun & Fougeron, 2000, 2002; Post, 2000) which cannot be discussed in full detail here.
In VF contexts, the auxiliary *hat* was realized with a falling nuclear pitch accent in 91% of the cases (76 times as H*L-, two of which were downstepped !H*L-, and 10 times as H+L* or H+!H*). These cases were followed by unaccented object nouns (88%, N=63) and unaccented non-finite verbs (92.5%, N=62). There were only four utterances with an unaccented *hat*: 2 times it was followed by an accented object noun (once as H*L- and once as H+L*) and twice by accented non-finite verbs (as H*L-). In Table 2 accent types realized on *hat* across contexts are illustrated:

<table>
<thead>
<tr>
<th>Accent types on hat</th>
<th>non-VF (%)</th>
<th>VF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H* L-</td>
<td>2.0</td>
<td>76.1</td>
</tr>
<tr>
<td>H+L* / H+!H*</td>
<td>0.0 / 0.0</td>
<td>4.5 / 10.4</td>
</tr>
<tr>
<td>Unaccented</td>
<td>98.0</td>
<td>9.0</td>
</tr>
</tbody>
</table>

Table 2: Distribution in percentages of accent types on the function word *hat*.

Not surprisingly, a chi-square test revealed a significant difference of use of accent types on the auxiliary *hat* across contexts: there were significantly more pitch accents on the auxiliary in VF contexts than in non-VF contexts ($X^2(2) = 88.5$, $p < .0001$, Cramer’s $V = 0.87^{9}$).

In Table 3 the accent types realized on the object noun across contexts are illustrated:

---

8 According to Rathcke and Harrington (2006) H+!H* and H+L* represent the same category.
9 For degree of freedom equal to 2, a V-value larger than 0.39 indicates a significant effect (Gravetter & Wallnau, 2006).
<table>
<thead>
<tr>
<th>Accent types on the object noun</th>
<th>non-VF (%)</th>
<th>VF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H* L-</td>
<td>64.0</td>
<td>4.5</td>
</tr>
<tr>
<td>H+L*</td>
<td>25.5</td>
<td>1.5</td>
</tr>
<tr>
<td>L*+H / L+H* L-</td>
<td>2.1 / 6.3</td>
<td>0.0 / 0.0</td>
</tr>
<tr>
<td>Unaccented</td>
<td>2.1</td>
<td>94.0</td>
</tr>
</tbody>
</table>

Table 3: Distribution in percentages of accent types on the object noun.

In non-VF contexts German speakers were more likely to produce a pitch accent on the noun object than in VF contexts (X² (3) = 92.5, p < .0001, Cramer’s V = 0.89).  

Finally, Table 4 displays the accentual realizations realized on the non-finite verb in both contexts.

<table>
<thead>
<tr>
<th>Accent types on the non-finite verb</th>
<th>non-VF (%)</th>
<th>VF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H* L-</td>
<td>0.0</td>
<td>7.5</td>
</tr>
<tr>
<td>H+L*</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Unaccented</td>
<td>98.0</td>
<td>92.5</td>
</tr>
</tbody>
</table>

Table 4: Distribution in percentages of accent types on the non-finite verb.

The difference in accent distribution on the non-finite verb across pragmatic contexts approached significance; however, the Cramer’s test showed no strong association between distribution of accent and context (X² (2) = 5.01, p = 0.08, Cramer’s V = 0.21, for an interpretation of Cramer's V see footnote 9). Post-hoc analyses of the data revealed that the pitch accent H* L- and H+L* only occurred in utterances in which the number of syllables following the auxiliary contained 9 and 10 syllables (see Appendix A for a list of the elicited target utterances). We therefore assume that the non-finite verb was accented for rhythmic reasons, avoiding a long stretch of
unaccented post-focal syllables (cf. for discussions on the Rhythm Rule, see, for instance, Liberman & Prince, 1977; Nespor & Vogel, 1989, 2007; Selkirk, 1984). Figures 2(a)-(b) show examples of the most typical contours realized by our German speakers in VF and in non-VF condition.

Figures 2(a)-(b): Example pitch tracks of non-VF utterance (upper panel) with an unaccented auxiliary hat and an accented object noun die Tür and of VF utterance (lower panel) with an accented auxiliary hat and an unaccented object noun den Reifen, spoken by a female German speaker. In all the figures, the pitch range is shown linearly from 150 to 300 Hz.

We will now turn to the occurrence and phonological form of hat patterns. In only 12% of the VF cases (N=8), speakers realized a hat pattern contour, i.e. a pitch rise on

\(^{10}\) For degree of freedom equal to 3 a V-value larger than 0.29 indicates a significant effect (Gravetter & Wallnau, 2006).
the possessive pronoun of the utterance-initial prepositional phrase *Auf meinem Bild* (“In my picture”) (all cases were realized with a L*+H, see Appendix B), a sustained high pitch (not changing more than 10 Hz) and a pitch fall either on the auxiliary (twice H+L*; twice H*L-) or on the object noun (once H+L* and twice H*L-). There was one case in which the pitch fall occurred only on the non-finite verb (H* L-). One example of hat pattern realized as H+L* is shown in Figures 3.

![Example pitch tracks of hat pattern with a pitch fall on hat (realized as H+L*).](image)

3.1.2 Discussion

In German, the intonational realization of both the non-VF and VF utterances were very systematic and – at least as accent placement is concerned – fairly consistent with our predictions derived from previous studies. Both conditions were disambiguated already at the auxiliary. As expected, in non-VF contexts, the auxiliary was generally unaccented and the nuclear accent was produced on the object noun, which is assumed to be the default location for broad focus cases (Büring, 2006; Uhmann, 1991, among others). In VF cases, the falling nuclear accent (H* L-) was realized on the auxiliary *hat*, optionally followed by a nuclear accent with an early peak (H+L* or H+!H*). In the majority of the cases the post-focal elements (object
noun, and non-finite verb) were deaccented. This is typical for many languages including German (Féry & Kügler, 2008; Hanssen, Peters, & Gussenhoven, 2008; Xu & Xu, 2005).

In a few cases only (12% of the VF cases), German speakers produced the hat pattern contour, which has been associated with double contrast in the literature (Büring, 2003; Mehlhorn, 2001). Hat patterns were always produced with an $L^*+H$ prenuclear accent but the nuclear accent was produced either with a $H^* L$- or a $H+L^*$. They were realized only in VF contexts thus confirming the contrastive nature of this prosodic contour in German. However, our findings show that the hat pattern contour is not the prototypical contour to signal a double contrast in German (for different interpretations of the hat pattern in German, Ambrazaitis & Niebuhr, 2008; Féry, 1993). Furthermore, our results suggest that the characteristic part of the hat pattern is the prenuclear accent (realized as $L^*+H$) and the high transition between the accents, rather than the identity of the nuclear accent.

3.2 French data

The interrater reliability for the raters was 95.3% for the presence or absence of initial accents on the auxiliary or the non-finite verb (Kappa = 0.93) and 93.0% for the object noun (Kappa = 0.79$^{11}$ for HiLL%, LHiL% and Unaccented).

3.2.1 Results

Given previous studies on the effect of focus on phrasing (Chen & Destrue, 2010; Dohen & Loevenbruck, 2004; Féry, 2001) as well as presence of an intermediate phrase level between the subject noun and the verb phrase in broad focus SVO

---

$^{11}$ Values of kappa above 0.60 show good to excellent agreement among different raters and values of 0.40 or less show fair to poor agreement (Gravetter & Wallnau, 2006).
structures (D’Imperio & Michelas, 2010; Michelas & D’Imperio, 2010), we first looked for presence of phrase breaks realized between the subject noun and the verb phrase across the two pragmatic contexts. Phrase breaks were coded in cases, in which there was a clear pause or in cases, in which the auxiliary was glottalized. Glottalization at the onset of the vowel was determined visually from the acoustic signal and coded when there were either irregular glottal vibrations in the initial part of the vowel or isolated glottal pulses before it (Dilley, Shattuck-Hufnagel, & Ostendorf, 1996). Phrase breaks were realized in 50% of non-VF cases (17 out of 34 cases) and in 45% of the VF cases (31 out of 69), ($X^2(1) = 0.08, p = 0.8$).

On the basis of previous studies reporting the occurrence of initial accents on the left edge of focused APs (German & D'Imperio, 2010), we then looked for the presence of initial accents realized on the left edge of the verb phrase in both contexts. Verb phrases were made of a function word (i.e. the auxiliary) followed by a disyllabic or a trisyllabic content word (i.e. the non-finite verb). For the identification of initial accents, we largely followed the criteria defined by German and D’Imperio (2010). More specifically, initial accents realized on the auxiliary were coded as FW-Hi, those on the first two syllables of the content word as CW-Hi, and the absence of an initial accent as No-Hi. Initial accents realized on the content word (CW-Hi) were all characterized by a peak on one of its first two syllables always preceded by an elbow (L)$^{12}$ between the function word and the content word. This initial accent was either followed by a dip (L) preceding the final accent (LHiLH*), or by a plateau extending from one of the first two syllables of the content words until the end of the final accent (LHiH*), or by a low tonal pitch accent (LHiL*) or by no final pitch accent at all (LHi). Such CW-Hi were more frequently realized without a preceding

---

$^{12}$ Initial and/or late elbows were identified semi-automatically and hand-corrected by using the procedure described in Welby (2006, p.351).
phrase break (70%, \(N=29\)) than with a phrase break (30%, \(N=12\)) \(\chi^2(1) = 7.05, p < 0.01, \text{Cramer's V}= 0.28\). In verb phrases with disyllabic content words (i.e. non-finite verbs such as \textit{vidé}) the initial accents were located on the first syllable of the non-finite verb (e.g., \textit{vidé}). In trisyllabic non-finite verbs (e.g., \textit{réveillé}), the initial accents were either located on the first \((N=16)\) or on the second syllable \((N=12)\).

Initial accents realized on the auxiliary (FW-Hi) were preceded by a phrase break in 83\% of the cases \((N=19, \chi^2(1) = 9.783, p < 0.001, \text{Cramer's V} = 0.387)\). When the initial accent was preceded by a phrase break its peak was either upstepped \((N=11)\), downstepped \((N=5)\) or produced with the same pitch (high plateau) as the final accent in the preceding AP \((N=3)\). Initial accents not preceded by a phrase break \((N=4)\) were always realized as a high plateau, starting from the final accent of the preceding AP and extending over the auxiliary\(^{14}\). The initial accent on the auxiliary was either followed by a final accent (HiLH*) or by a low tonal target (HiLL*) both realized on the last syllable of the non-finite verb. Hence, there was no case, in which a phrase break was realized after the initial accent on the auxiliary. In other words, initial accents were always phrased in the same AP together with the following content word (i.e., the non-finite verb). In Figure 4 a stylized form of initial accents realized on the FW as HiLH* and HiLL* (the low target is indicated by the dotted line) is illustrated.

\(^{13}\) For degree of freedom equal to 1 a V-value between 0.30 and 0.50 indicates a medium effect. A V-value between 0.10 and 0.30 indicates a small effect, and above 0.50 a large effect.

\(^{14}\) The average semitone difference between the peak on the auxiliary and the peak on the subject noun was 2.15 semitones for upstepped rises \((SD = 0.23)\), and of -1.45 semitones \((SD = 0.93)\) for downstepped rises. The initial rises that were not preceded by a phrase break were characterized by an average of 3.4 Hz difference between the beginning and the end of the plateau.
Figure 4: Stylized form of initial accent realized on the FW and followed by a final accent (H\*) or by a low tonal target (L\*) as indicated by the dotted lines.

Table 5 displays the distribution of initial accents on the auxiliary (FW-Hi), on the first two syllables of the content word (CW-Hi) vs. cases, in which the initial accent was not realized at all, in VF contexts compared to non-VF contexts.

<table>
<thead>
<tr>
<th>Initial Accents (Hi)</th>
<th>non-VF (%)</th>
<th>VF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FW-Hi</td>
<td>0.0</td>
<td>33.3</td>
</tr>
<tr>
<td>CW-Hi</td>
<td>35.3</td>
<td>42.0</td>
</tr>
<tr>
<td>No-Hi</td>
<td>64.7</td>
<td>24.7</td>
</tr>
</tbody>
</table>

Table 5: Distribution in percentages of initial accents realized on the verb phrase

As we can see from the table, initial accents on the FW were realized only in VF cases. We then tested whether the above reported effect of context was reliable over and above speaker specific preferences and number of syllables of the verb phrase. To this end, we conducted a binomial logistic regression model (Baayen, 2008; Pinheiro & Bates, 2002) with FW-Hi as dependent variable (the first row of the Table 5 was coded as ‘Yes’ and the remaining rows coded as ‘No’), and context as predictor. We further added the number of syllables contained in the target AP (i.e. the verb phrase) as predictor as it might influence the occurrence of initial accents due to rhythmic reasons (e.g., Jun & Fougeron, 2002, cf. section 1.4). Finally, speaker and item were added as crossed random factors. The model revealed that there was a significant

---

15 Since binomial logistic regression model cannot be calculated if there are no instances in a given condition, we replaced one
effect of context \((z = 2.98, p < 0.001)\) but no effect of number of syllables \((z = 0.25, p = 0.80)\) and no interaction \((z = 0.91, p = 0.35)\). There were significantly more initial accents on the FW in VF than in non-VF contexts. This implies that there is a categorical distinction (use of different accent patterns) between the two contexts. In Table 6 we provide more detailed information on the accentual realizations of the whole verb phrase and their frequency of occurrence in the two contexts.

<table>
<thead>
<tr>
<th>Accent patterns on the verb phrase</th>
<th>non-VF (%)</th>
<th>VF (%)</th>
<th>Stylized forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>HiLH* / HiLL*</td>
<td>0.0 / 0.0</td>
<td>27.6 / 5.8</td>
<td>A REVEILLE A REVEILLE</td>
</tr>
<tr>
<td>LHiL* / LHi</td>
<td>0.0 / 0.0</td>
<td>11.6 / 4.3</td>
<td>A REVEILLE A REVEILLE</td>
</tr>
<tr>
<td>LHiH* / LHiLH*</td>
<td>29.4 / 6.0</td>
<td>17.4 / 8.7</td>
<td>A REVEILLE A REVEILLE</td>
</tr>
<tr>
<td>LH* / LLH*</td>
<td>41.1 / 23.5</td>
<td>20.3 / 4.3</td>
<td>A REVEILLE A REVEILLE</td>
</tr>
</tbody>
</table>

Table 6: Distribution in percentages of accent patterns on the verb phrase and relative stylized forms aligned with an example verb phrase *a réveillé*.

Table 6 shows that there are four accent patterns HiLH*/HiLL* and LHiL*/LHi that were only realized in VF contexts (in 50% of the cases). In Figures 5(a)–(b), typical patterns (HiLH* and LHiL*) that only occurred in VF contexts are illustrated.

instance without initial rise on the FW with a “Yes” in the non-VF condition (Braun & Chen, 2010, footnote 3).
Figure 5(a)-(b): Example pitch tracks of VF utterances with initial accent on the auxiliary (upper panel) realized as HiLH* and initial accent on the content word followed by a low tonal target (lower panel) realized as LHiL*, spoken by two female French speakers. Phrase breaks are indicated by ‘#’.

Hence in almost half of the cases (49.3%) French speakers phonologically distinguished VF from non-VF cases in the verb phrase, i.e. the accent patterns HiLH*/HiLL*/LHiL*/LHi were only used in VF contexts. This is achieved in two ways: 1) either by placing an emphatic initial accent on the auxiliary followed by a final rising accent (HiLH*) or by a low tonal target (HiLL*); 2) or by placing an emphatic initial rise on the CW followed by a low tonal target (LHiL*) or by no final accent (LHi).

In the remaining cases (50.7%), French speakers realized the accent patterns LHiH*/LHiLH*/LH*/LLH* equally often in both contexts (LLH*/LH*: N=22 in non-VF vs. N=17 in VF; LHiH*/LHiLH*: N=12 in non-VF vs. N=18 in VF). Importantly, in accentual realizations with both an initial and a final accent (LHiH*/LHiLH*: N=3/9 in non-VF vs. N=9/9 in VF), the initial accent sounded more prominent than the final accent. This auditory impression was also supported by an acoustic analysis of the two peak heights of the accents across the two contexts. As above mentioned, on the basis of previous study (Dahan & Bernard, 1996; Dohen & Loevenbruck, 2004; Jun & Fougeron, 2000, among others), we compared the acoustic realization of initial
accents in the two conditions, i.e. for accent patterns containing initial accents on the CW (LHiLH*/LHiH*), we compared the height of the peak of the initial accent on the CW (H1 in semitones, cf., Nolan, 2003) to the height of the peak of the final accent of the last syllable of the non-finite verb (H2 in semitones) and calculated the semitone differences between the two peaks. The semitone difference (ΔST) was calculated with the formula: ΔST = 12*log₂ (H1/H2). We tested whether the F0-peak of the initial accents (H1, in semitones, cf., Nolan, 2003) was higher than the peak of the final accents (H2, in semitones) in VF contexts compared to non-VF contexts. With the same procedure as above, we ran a linear mixed effect model (Baayen, 2008) with tonal scaling as dependent variable and accent pattern (LHiH*/LHiLH*) and context (non-VF vs. VF) as fixed factors. We further included gender as a fixed factor but removed it again as it was not significant (p = 0.7). Results showed that there was a significant effect of context (p < 0.001) on the scaling difference but not of accent pattern (p=0.9): in VF contexts the tonal scaling difference between the initial accent and the final accent decreases significantly with respect to non-VF contexts: both accents sound equally prominent in VF contexts, whereas the final accent has a higher peak than the initial in non-VF contexts. This difference is illustrated in Figure 6:

---

16 The HiLH* accent pattern could not be included as it only occurred in VF context. For this accent type also, the initial accent had a higher peak than the final accent (average semitone difference = 2.3, SD = 1.2, p < 0.05).
Figure 6: Barplots of LH\textsubscript{i}LH* and LH\textsubscript{i}H* in VF and non-VF contexts. Mean values are based on the statistical model and whiskers represent standard error. Large values indicate that the final accent is higher than the initial accent, zero indicates that the two accents are equally high, negative values indicate that the initial accent is higher than the final accent.

Similar phonological contours realized in VF and non-VF cases with tonal scaling differences are shown in Figures 7(a)–(b).
Figure 7(a)-(b): Example pitch tracks of non-VF (upper panel) and VF (lower panel) spoken by a female French speaker showing tonal scaling differences between the two peaks of the two accents. In the VF example the initial rise is as higher as the final rise.

Finally, given that post-focal APs in French are usually unaccented in contrastive contexts, a significant difference in the tonal realizations of the object noun was found across the two contexts. There were significantly more unaccented object nouns in VF than in non-VF contexts as shown in the Table 7 ($X^2 (2) = 74.6, p < .0001$, Cramer’s $V = 0.85$). This difference in post-focal realization is also shown in Figures 7(a)-(b).

<table>
<thead>
<tr>
<th>Accent patterns on the object noun</th>
<th>non-VF (%)</th>
<th>VF (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HiLL%</td>
<td>41.2</td>
<td>2.9</td>
</tr>
<tr>
<td>LHiL%</td>
<td>52.9</td>
<td>4.3</td>
</tr>
<tr>
<td>Unaccented</td>
<td>5.9</td>
<td>92.8</td>
</tr>
</tbody>
</table>

Table 7: Distribution in percentages of accent patterns on the object noun

3.2.2 Discussion

In half of the cases, French participants marked Verum Focus by using categorically distinct accent types compared to non-VF conditions. More specifically, the verb phrase was realized with an emphatic initial accent on the auxiliary (accent patterns HiLH*/HiLL*, 33% of the cases) or on the first syllables of the non-finite verb without an accompanying final rising accent (LHiL*/LHi, 16% of the cases). On the other hand, initial accents on the first syllables of the content word were also produced frequently in non-VF contexts, but they were always followed by a final accent (LHiH*/LHiLH*). In the other half of the Verum Focus cases (50.7%), accentual realizations (LHiLH*/LHiH*/LH*/LLH*) were produced in both contexts equally frequently. However, acoustic analyses revealed tonal scaling differences of the peak of the initial and final accents across contexts. More specifically, in VF
contexts the two accents sounded equally prominent.

Our attention particularly goes to the initial accents realized on the function words (HiLH*/HiLL*). In most of the cases they were all realized with an emphatic initial accent and preceded (but never followed) by a phrase break. If we compare our results on the initial accent realized on FW-Hi with the accent types reported by Jun and Fougeron (2002) and Welby (2006), we notice that the accent HiLH* was only rarely found by Jun and Fougeron (2002) and Welby (2006) in their datasets. The authors attributed the scarce presence of HiLH* to the fact that this accent pattern is typically realized with APs made of content words only. By contrast, in our data we noticed that HiLH* is specifically favored by speakers when they want to mark Verum Focus on the auxiliary. The Hi tone is then realized on the function word a and not on the content word of the AP. Hence, the use of certain accent types might not be just a question of word type (function vs. content word), but also of pragmatic context differences.

The analysis of the current French data reveals that the use of accent types on the verb phrase can partly be determined by the pragmatic context (i.e. HiLH*/HiLL* only occurred in VF contexts). More importantly, our data showed that monosyllabic function words (in our case auxiliaries) can be accented (with an initial accent) in contrastive environments (Verum Focus). There were no instances of initial accents on the auxiliary in non-VF cases even though the context might have theoretically licensed their occurrence as in both VF and non-VF, the focus domain starts with the auxiliary. Our data hence suggest that French speakers differentiate broad focus (here: non-VF) from narrow focus contexts (here: VF).

Our results further reveal that in non-VF cases, French speakers either use accentual patterns with an initial accent on the first syllables of the content word...
(LHiLH*/LHiH*), followed by a final accent, or realized with no initial accent at all (LLH*/LH*). In the former case there are, however, clear phonetic differences with respect to the Verum-Focus realizations, the final accent sounds more prominent than the initial one, whereas in VF cases the semitone difference between the two accents decreases. Interestingly, the gradient differences in peak scaling of the initial and final accents make the final accent less salient than the initial accents and hence go in the same direction as the categorical differences reported before.

4. Conclusions

In the present study, the prosodic status of function words (i.e. a monosyllabic auxiliary) was tested by comparing semi-spontaneous Verum Focus utterances with non-Verum (predicate) Focus utterances in German and French.

First of all we investigated whether the auxiliary can get accented in a language, German, where function and content words alike can be pitch accented. We found that in German the auxiliary received a nuclear pitch accent (Baumann et al., 2007) in the intended context (Verum Focus) thus confirming the hypothesis that its prosodic status in a contrastive environment does not differ from content words. When function words are focused their prominence pattern changes (from unaccented in non-VF focus contexts to nuclear accented in VF contexts) and they become stressed and carry a pitch accent. More importantly, this pattern was fairly consistent across speakers. There were also some occurrences of hat patterns produced with a rising accent on the possessive pronoun and a falling accent on the auxiliary in Verum Focus contexts. Although they represented the minority of the cases, we confirmed their rare presence for semi-spontaneous speech (cf. Braun, 2006) and added evidence to previous studies on the contrastive nature of such a structure in German. This first
part of the investigation therefore adds experimental evidence to a phenomenon, Verum Focus, which has been addressed in the Germanic literature mainly from a semantic viewpoint (Dimroth, 2004; Féry, 2006; Krifka, 1999; Reis & Rosengren, 1997; Sudhoff, 2010, among others). Also, it confirms previous studies on how function words can carry distinctive pragmatic meanings along with their accented status (Féry, 2006; Krifka, 1999; Reis & Rosengren, 1997; Sudhoff, 2010, among others). Furthermore, we collected new data on the phonological form of hat patterns, elicited in a more spontaneous setting.

Given the existence of cross-linguistic differences in the interface between intonation and information structure, we investigated how Verum Focus is realized in French. As discussed, French is a language which poses interesting questions about the prosodic status of function words in phrase-initial position (Delais-Roussarie, 1995; 1999, among others). It was shown that French speakers were able to intonationally express pragmatic context differences both categorically (use of different accent types across pragmatic contexts) and gradiently (tonal scaling differences for accent patterns that were realized in both pragmatic contexts). More specifically, we found that one third of the cases were realized with an initial accent on the focused monosyllabic function word. This finding confirms previous studies on French focus marking (Clech-Darbon et al., 1999; Di Cristo, 1998; Rossi, 1985; Touati, 1987) and more recent studies on the pragmatic function of initial accents (Astésano et al., 2007; German & D'Imperio, 2010). However, unlike previous studies which have investigated the intonational realization on focused content words (Féry, 2001; Jun & Fougeron, 2002, among others), mostly in read speech, in the present investigation we provided evidence from semi-spontaneous data and showed that French speakers can focus monosyllabic phrase-initial function words. This finding
therefore opens a new link between prosodic status of function words and information structure in French.

Yet, we clearly see cross-linguistic differences in the prosodic status of function words across the two languages. While in German the focused auxiliary was promoted to have a nuclear pitch accent, in French the initial accent associated to the function word was constantly accompanied by the final accent on the final syllable of the content word (as H or as L tone). Also, the function word was never produced in an isolated AP (with no right-alignment of the AP-boundary) but always phrased together with the content word. Finally, the presence of the initial accent on the auxiliary was optional. This confirms Jun and Fougeron's (2002) and Welby’s view (2006) on the nature of the initial accent as a phrase accent rather than as a pitch accent. Nevertheless, we noticed the pragmatic function played by the initial accent in Verum Focus contexts compared to non-VF ones. This is supported by findings on tonal scaling difference between the two accents and presence of a phrase break (Dahan & Bernard, 1996; Séguinot, 1976) correlated with occurrence of initial accent on the function word. A perception study will follow in order to test to what extent the prominence of the initial accent compared to the final accent in the different accentual realizations (HiLH*/LHiH*/LHiLH*) plays a role in disambiguating Verum Focus from non-Verum Focus contexts is in preparation.

The large variation of intonational patterns in French compared to German can be accounted for by cross-linguistic differences (in addition to speaker ‘preferences’) in information structure marking (Dimroth et al., 2010). As discussed above, Dimroth’s study revealed that Romance speakers (French and Italian) did not use linguistic means (i.e. Verum Focus, particles) specifically highlighting the shift from a negative antecedent sentence into a positive sentence to the same degree as speakers
of Germanic languages (German and Dutch). Rather, other linguistic means highlighting the shift from one topic to another (i.e. contrastive strong pronouns in French, cleft sentence in Italian) were preferred by Romance speakers in order to convey such a contrast. The preference for other linguistic means on the one hand and the lack of an inventory of focus particles on the other, might discourage French speakers to systematically mark Verum Focus via intonation. This could partly explain why French speakers produced a wide range of intonational realizations when confronted with controlled tasks such as the one in the present study where speakers are highly encouraged to mark a shift only on the polarity with no possibility of word order change. There is one thing that might speak in favor of this hypothesis: as said above (cf. section 1.5 and 2.2) the experimental set-up did not encourage explicit corrections but rather a general form of contrast where participants were given the option to either simply report what they saw on their pictures or explicitly highlight the difference between their picture and the confederate speaker’s picture. Nevertheless, we noticed that despite the free choice, Germans were consistent in marking the function word *hat* (and therefore the VF), while French speakers were not.

In future work we will test whether the variability we found in French Verum Focus is also present in disyllabic function words (e.g., *avaient* vs. *a*), in copula constructions (e.g., *est heureuse*, “is happy”) and in finite lexical verbs (*mange*, “eats”) to be able to give a more concise approach to the prosodic status of function words in French. A second goal is to investigate how learners of these languages are able to acquire focus marking and how listeners from each language, as well as learners, perceive and disambiguate these structures from non-Verum Focus contexts.
References


LEONETTI, M., & ESCANDELL-VIDAL, M. V. (2009). Fronting and verum focus in Spanish. In A. Dufter & D. Jacob (Eds.), Focus and Background in Romance Languages (pp. 155-204). Amsterdam: John Benjamins.


Appendix A: German and French elicited utterances of the dialogue-game

Table A1. German materials including information on the number of syllables of the non-finite verb.

<table>
<thead>
<tr>
<th>number of syllables of non-finite verb</th>
<th>Elicited utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Auf meinem Bild hat der Waldarbeiter den Baum <em>gefällt</em>.</td>
</tr>
<tr>
<td>2</td>
<td>Auf meinem Bild hat der Straßenkehrer den Gehsteig <em>gefegt</em>.</td>
</tr>
<tr>
<td>2</td>
<td>Auf meinem Bild hat die Frau die Blume <em>gepflückt</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat der Jaguar die Schwalbe <em>gebissen</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat der Fleischer das Fleisch <em>geschnitten</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat der Herr die Krawatte <em>geknotet</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat der Polizist die Tür <em>aufgebrochen</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat der Herr die Sektflasche <em>geöffnet</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat der Junge den Reifen <em>zerstochen</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bier hat der Obdachlose das Bier <em>getrunken</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat das Kind die Bonbons <em>gegessen</em>.</td>
</tr>
<tr>
<td>3</td>
<td>Auf meinem Bild hat das Mädchen den Geldschein <em>zerrissen</em>.</td>
</tr>
</tbody>
</table>

Table A2. French materials including information on the number of syllables of the non-finite verb.

<table>
<thead>
<tr>
<th>number of syllables of the non-finite verb</th>
<th>Elicited utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>sur mon image la dame a <em>cueilli</em> la tulipe</td>
</tr>
<tr>
<td>2</td>
<td>sur mon image le garçon a <em>crevé</em> le ballon</td>
</tr>
<tr>
<td>2</td>
<td>sur mon image le gorille a <em>avalé</em> la bague</td>
</tr>
<tr>
<td>2</td>
<td>sur mon image l'enfant a <em>mangé</em> le bonbons</td>
</tr>
<tr>
<td>2</td>
<td>sur mon image le renard a <em>vidé</em> le sac-à-dos</td>
</tr>
<tr>
<td>3</td>
<td>sur mon image l'enfant a <em>déchiré</em> le billet</td>
</tr>
<tr>
<td>3</td>
<td>sur mon image le jaguar a <em>attrapé</em> l'hirondelle</td>
</tr>
<tr>
<td>3</td>
<td>sur mon image le gardian a <em>arrêté</em> le ballon</td>
</tr>
<tr>
<td>3</td>
<td>sur mon image le bûcheron a <em>abattu</em> l'arbre</td>
</tr>
</tbody>
</table>
sur mon image l’éléphant a défoncé le parquet
sur mon image l’oiseau a réveillé le policier
sur mon image le balayeur a nettoyé le trottoir

Acknowledgements
This research is part of a PhD project (carried out by the first author) framed in the ANR-DFG project ‘LANGACROSS’ (DFG Grant and the Geschäftszeichen DI 808/1-1) awarded to Christine Dimroth. We would like to thank Tilman Harpe for help with the pictures, and Johan Weustink for programming the experiments. Furthermore, we are deeply indebted to Mariapaola D’Imperio, Bob Ladd and Barbara Hemforth for very insightful comments on a previous version of this paper.