JOURNAL ARTICLE

Foxes, deer, and *hedgehogs*: The recall of focus alternatives in Vietnamese

Annika Tjuka[1,2], Huong Thi Thu Nguyen[1] and Katharina Spalek[1]

[1] Humboldt-Universität zu Berlin, Berlin, DE

[2] Max-Planck-Institute for the Science of Human History, Jena, DE

Corresponding author: Annika Tjuka (tjuka@shh.mpg.de)

In tonal languages, the role of intonation in information-structuring has yet to be fully investigated. Intuitively, one would expect intonation to play only a small role in expressing communicative functions. However, experimental studies with Vietnamese native speakers show that intonation contours vary across different contexts and are used to mark certain types of information, for example, focus (Jannedy, 2007). In non-tonal languages (e.g., English), the marking of focus by intonation can influence the processing of focus alternatives (Fraundorf, Watson, & Benjamin, 2010). If Vietnamese also uses intonation to mark focus, the question arises whether the behavioral consequences of prosodic focus marking in Vietnamese are comparable to languages such as English or German. To test this, we replicate a study on memory for focus alternatives, originally carried out in German (Koch & Spalek, in progress), with Vietnamese language stimuli. In the original study, memory for focus alternatives was improved in a delayed recall task for focused elements produced with contrastive intonation in female speakers. Here, we replicate this finding with Northern Vietnamese native speakers: Contrastive intonation seems to improve later recall for focus alternatives in Northern Vietnamese, but only for female participants, in line with the findings by Koch and Spalek (in progress). These results indicate that prosodic focus marking in Vietnamese makes alternatives to the focused element more salient.

Keywords: Alternative semantics; information structure; delayed recall; memory; focus; intonation; Vietnamese

1. Introduction

The information conveyed by a given sentence is expressed differently across languages. For example, *novel* information might be highlighted by a change in word order (syntax), the addition of a specialized marker (morphology), or a variation of the intonation contour (phonology). These features can also be used to indicate *contrastive* information. The marking of both types of information is summarized by the term *focus*. If focus marking is present, it establishes a set of alternatives which are important for interpreting the meaning of an utterance (Krifka, 2006).

Building on the theoretical assumptions made by Krifka (2006), experimental studies provide empirical evidence for the relevance of focus alternatives in language processing (e.g., Braun & Tagliapietra, 2010; Fraundorf et al., 2010; Gotzner, Wartenburger, & Spalek, 2016; Husband & Ferreira, 2016). These studies have investigated the influence of intonation focus or focus-sensitive particles in English, German, and Dutch. As of yet, studies investigating the processing of focus alternatives in other languages are scarce (but see Yan & Calhoun, 2019, for data from Mandarin Chinese).

It has been assumed, for example by Lee-Wong (1994), that a language with a complex tone system and a variety of particles, like Vietnamese, would not use intonation contour to express pragmatic functions. However, studies based on natural speech corpora and experiments indicate that Vietnamese does use intonation to emphasize certain information (Hạ, 2012; Jannedy, 2007; see Michaud and Brunelle, 2016, for a comparative overview). Moreover, Jannedy (2007) showed that focus can be marked with intonation in the Northern Vietnamese dialect.

In our study, we investigate whether contrastive intonation improves later recall for focus alternatives in (Northern) Vietnamese. Before turning to the experiment, we will give some information on the Vietnamese language, discuss studies examining information structure in Vietnamese, and studies investigating the effect of intonational focus on memory in German and English. Based on this background, we expect that intonational focus establishes a set of alternatives in (Northern) Vietnamese just as it does in Germanic languages. Due to their increased salience, these alternatives will be remembered better in a delayed recall task.

1.1. The Vietnamese language

Vietnamese, an Austro-Asiatic language of the Mon-Khmer branch, is a tone language. It is now spoken by about 95 million people in Vietnam and by about four million speakers living elsewhere in the world. Tone has been used to classify three major dialect groups in Vietnam: North Vietnamese, South Vietnamese, and Central Vietnamese dialects (Hoàng, 1989; Vũ, 1982). Vietnamese is an isolating language with no marking of case, number, or gender and the majority of semantic units are single morphemes (Campbell, 2003). Word order in Vietnamese is S-V-O, with very few exceptions.

1.1.1. The Vietnamese tone system

The Vietnamese tone system is not consistent across different dialect groups in Vietnam. We will focus here on the tone system of Northern Vietnamese, since this is the variety we will be looking at. Northern Vietnamese tones are characterized by a combination of pitch height and voice quality. Michaud and Vũ (2004) and Brunelle (2009) (see also, Brunelle, Nguyễn, & Nguyễn, 2010) describe the eight tones of Northern Vietnamese as follows: *ngang* (A1) is high level with modal phonation; *huyền* (A2) is low falling with (usually) modal phonation; *sắc1* (B1) starts at mid pitch range and then rises quickly, its phonation is modal; *nặng1* (B2) starts in the pitch mid-range and falls dramatically, it also ends with a strong glottalization; *hỏi* (C1) falls strongly until it reaches a turning point, it ends with a laryngealization (breathy voice); *ngã* (C2) starts on a fall, is strongly laryngealized and rises strongly. Tones *sắc2* (D1) and *nặng2* (D2) appear in stop-closed syllables.

1.1.2. Intonation contours

The complex lexical tone system in addition to the variety of particles seems to provide very little space for (and necessity of) prosodic intonation cues for pragmatic functions in Vietnamese. Nevertheless, in an early description of the language, Thompson (1965) observed that pitch is used to express communicative functions.[1]

In addition to communicative functions, intonation can also signal linguistic functions. Đỗ, Trần, and Boulakia (1998) examined the prosodic realization distinguishing different sentence types. Their stimulus set comprised simple declarative sentences that were transformed into interrogative and imperative sentences. The sentences were constructed

---

[1] If variance in intonation contours is described, it is mainly in the framework of superposition (Fujisaki, 1988), for example, in terms of F0 contours. However, the question of what kind of tone-intonation interaction model best describes the Vietnamese system still needs to be examined in detail. For now, we depend on individual studies which investigate whether discourse-related functions can be expressed by intonation.

with six words each bearing the same tone. Each sentence used one of the six Vietnamese tones; see example (1) for a declarative sentence with tone *huyền*.

(1)   Bà   Hòa   tìm   nhà bà   liền
        3.SG Hòa(name) find house 3.SG immediately
        "Hoa found her house immediately."   (Đỗ et al., 1998)

Furthermore, they used long sentences and expressive dialogues in a narrative context. In a sentence reading task, participants were recorded while they read the sentences out loud. The speakers came from different regions in Vietnam. The acoustic analysis of the produced sentences showed that different pitch contours are used to distinguish pragmatic functions, such as differences between sentence types (Đỗ et al., 1998). Apart from a variation in global F0, Đỗ et al. (1998) identified syllable length and intensity as additional parameters to differentiate between the sentence meanings. These findings support the hypothesis that intonation plays a role in marking linguistic functions.

   In a natural speech corpus of one-word utterances, Hạ (2012) compared discourse particles and function words, for instance, *hả* 'isn't it/aren't it (interrogative particle),' in citation form with their pitch contours in different pragmatic contexts. The corpus consisted of telephone conversations between Northern Vietnamese speakers. The citation forms were elicited by recording controlled dialogues which were added to the corpus. The analysis was based on the model of Autosegmental-Metrical Phonology (e.g., Goldsmith, 1976; Gussenhoven, 2004; Ladd, 2008). Her findings support the assumption that different intonation contours are characteristic for a specific pragmatic context. But most interesting is the result that in some cases, intonation contours completely overrode the lexical tone of the words in different pragmatic contexts, for example, backchannels or turn-yielding (Hạ, 2012). In other cases, intonation contours followed the lexical tone. The interplay between tone and intonation further supports the assumption that Vietnamese systematically uses intonation irrespective of its complex tone system. However, Hạ (2012) suggests that a pragmatic context can evoke a certain intonation contour but these strategies might not be grammaticalized which means that an intonation contour cannot be attributed to a specific pragmatic context.

   Brunelle, Hạ, and Grice (2012) analyzed intonation patterns of different speakers of the Northern Vietnamese dialect to investigate whether they are identical across participants. They compared the intonation of sentences that end with the particle *không*. This sentence-final particle is used to form segmentally and tonally identical interrogative and declarative sentences (Brunelle et al., 2012). The particle *không* bears different meanings: 'empty,' 'no,' and as a focus particle 'only.' The data was systematically controlled for segments, tones, and syntactic structures. Their results showed that different communicative functions are expressed with a variation in the three acoustic parameters F0, intensity, and duration. However, the different meanings of *không* did not influence the intonation of declarative sentences (Brunelle et al., 2012). In addition, the authors found a notable inter-speaker variation caused by speaker-specific strategies for marking communicative functions with differences in mean F0 values. Brunelle et al. (2012) conclude that sentence mood and attitudes are most commonly distinguished by sentence-final particles with ambiguous sentences being extremely rare. Therefore, intonation patterns that indicate a certain meaning of the particle *không* may not be fully grammaticalized in the Northern Vietnamese dialect and only enhance other strategies for marking communicative functions (Brunelle et al., 2012). Although the intonation contours for a specific function might vary across speakers, the studies showed that Vietnamese uses a set of prosodic indicators to distinguish between pragmatic contexts.

1.1.3. Focus marking

One of the pragmatic functions which can be realized with a shift in the intonation is focus. The marking of focus as an information structural phenomenon is used to indicate new or contrastive information in an utterance (Krifka, 2006). The latter function of focus is illustrated in the following examples. Capital letters indicate focus accent and the focused element is marked with the subscript F:

(2)   Tina found [the SOCKS]$_F$ under the bed.

(3)   [TINA]$_F$ found the socks under the bed.

Although the sentences are semantically and syntactically identical, they lead to different interpretations. Example (2) implies that Tina did not find other pieces of clothing, for example, t-shirts or scarves under the bed. In example (3), the speaker intends to indicate that it was Tina and not Lisa, Robert, or Kate who found the socks under the bed. In addition, focus frequently occurs in response to a question, for example, *Who found the socks?*. The part of the answer which corresponds to the *wh*-constituent is often expressed with a focus intonation.

Vietnamese uses a large number of particles to express grammatical features. Three focus-sensitive particles (*thậm chí* 'even,' *chỉ* 'only,' *cả* 'also') have been described by Hole (2008). However, just as their counterparts in English or German, these particles add extra meaning components to the interpretation of an utterance (e.g., exhaustivity). There is no particle that marks bare focus. Since word order variations are not used to express focus, it has to be marked in the intonation contour. A seminal study on the interplay of intonation focus and lexical tones has been carried out for Mandarin Chinese by Xu (1999). He found that focus expands the local tone contours for non-final focused words, whereas the pitch range was suppressed in post-focal words.

Studies on prosodic focus marking in Vietnamese are scarce. Jannedy (2007) conducted the first experimental study investigating this topic in Northern Vietnamese. The study looked at both the production of sentences in a contrastive focus condition and the interpretation of those sentences. The first part consisted of a reading task with a question-answer paradigm between two native Vietnamese speakers of the Northern dialect. The paradigm was constructed as a casual conversation and included question-answer combinations leading to focus on different parts of the response, for example, the verb (4-a) or subject (4-b):

(4)   a.   Phương   đi gì?   – Phương   đi [xe đạp]$_F$.
              Phương(name) go what – Phương(name) go CLF bicycle
              "What does Phương do? – Phương rides his bike."

       b.   Ai   đi xe   đạp?   – [Phương]$_F$ đi xe   đạp.
              who go CLF bicycle – Phương(name) go CLF bicycle
              "Who is riding a bike? – Phương rides his bike."   (Jannedy, 2007)

The acoustic analysis of the two speakers revealed that F0 maxima in subject- and verb-focus utterances appeared earlier than in the sentential-focus and object-focus sentences where the pitch excursions appeared towards the end of the utterances (Jannedy, 2007). A local change in F0 patterns in the different focus conditions was also present. The visualization affirms clear accentual prominence on the focused element for the subject- and verb-focus (Jannedy 2007). The analysis revealed differences in the overall duration of the utterance and the subject- and verb-constituents.

The second part of the study included a forced choice identification perception task of the recorded question-answers with six other Vietnamese native speakers of the Northern dialect. Participants heard the recorded answers auditorily and had to choose the correct question. The results showed that focus intonation increased the correct choice of a question. The findings from Jannedy (2007) that sentences with different intonation patterns were classified correctly by native speakers indicate that prosodic focus marking is used in Northern Vietnamese. In addition, the acoustic analysis demonstrated that focus seems to be marked predominantly through prosody, including F0 movement, duration, and intensity in Northern Vietnamese. However, similar to Brunelle et al. (2012), inter-speaker variation of mean F0 values was also found in this study.

Brunelle (2017) investigated whether word stress in disyllabic words is present in Southern Vietnamese, but he also in passing discussed acoustic properties of focus marking with contrastive information. The stimulus set of the study contained three different sentence lists. One of them was used to produce corrective focus. It included frame sentences with contextual prompts to elicit natural corrective focus, see (5) as an example.

(5)   a.   Nhà   ba   má   Tuấn   có   nhiều tủ.sách   đẹp.
              house father mother Tuấn(name) have many bookshelves beautiful
              "Tuan's parents' house has a lot of beautiful bookshelves."   (context)

       b.   Không.phải! Họ   có   nhiều bàn.ghế   đẹp.
              no   they have many furniture beautiful
              "No! They have a lot of beautiful FURNITURE."   (target)
                                                                          Brunelle (2017)

The acoustic analysis of the reading task revealed that the two syllables of words were longer under focus than in the non-focal condition (Brunelle, 2017). Furthermore, most speakers raised their F0 to mark corrective focus. However, Brunelle (2017) found tone-specific differences: Increase of F0 was mainly present in high tones (*sắc* and *ngang*) whereas the F0 of the falling tone (*huyền*) stayed low. These findings can only be considered as tendencies, as the data set is not sufficient for a generalization. In addition, only 3 out of 10 participants relied exclusively on syllabic lengthening (Brunelle, 2017). In contrast to the results in Jannedy (2007), intensity was not a determining factor in marking focus.

In a recent study by Miller, Athanasopoulou, Pincus, and Vogel (2015), participants were prompted to produce sentences with target words with either tone *sắc* or tone *ngã*. Depending on the question they had to respond to, focus was either on the critical word or not. Miller et al. (2015) observed no effects of focus on pitch and phonation; instead, focus was marked by changes in duration and spectral energy.

It seems to be the case that sentence comprehension processes underlie a language-universal mechanism, as shown by Ip and Cutler (2017). They investigated prosodic entrainment in Mandarin Chinese in a phoneme detection experiment. While listening to sentences, the participants had to identify the target sound [ph]. The results showed that intonation is used to forecast upcoming focus by speakers of a tonal language although pitch contours are already used for lexical tone perception (Ip & Cutler, 2017). Regardless of cross-linguistic differences in the production of focus, the authors demonstrated that listeners use preceding intonation cues to predict focus on the target item which is also found in Germanic languages (Ip & Cutler, 2017). Interestingly, they observed pitch cues such as greater F0 range expansion three or four syllables before the onset of the predicted accent to upcoming focused elements. These findings further support the assumption that despite the lexical tone system of a language, focus intonation is used to convey communicative functions.

1.2. Memory representations for focus alternatives

As already touched upon in the previous paragraphs, focus introduces a set of alternatives that are related to the focused element (see Rooth, 1992). The sentences in examples (2) and (3) (repeated in (6) and (7)) differ in terms of their focus alternative sets. In (6), the set consists of socks, t-shirts, scarves, …, whereas in (7) the set consists of individuals, for example, Tina, Lisa, Robert, Kate, ….

(6)   Tina found [the SOCKS]$_F$ under the bed.

(7)   [TINA]$_F$ found the socks under the bed.

In recent years, the processing of focus alternatives was studied in psycholinguistic experiments with English, Dutch, and German speakers (e.g., Braun & Tagliapietra, 2010; Fraundorf et al., 2010; Gotzner et al., 2016). One research question concerned whether the presence of focus can facilitate the memory for words belonging to the focus alternative set. This was either investigated with intonational focus marking (Fraundorf et al., 2010) or focus particles (Spalek, Gotzner, & Wartenburger, 2014).

   The study by Fraundorf et al. (2010) examined the processing of pitch accenting and its effect on memory for focus alternatives. Participants listened to auditory stimuli consisting of short discourses in English. First, a context passage (8) introduced the two focus alternatives. The sentence was followed by a continuation passage (9) in which one word from each contrast set was repeated. The authors manipulated whether the chosen elements from the contrast set in (9) were spoken with an $L+H^*$ accent or with an H^* accent.

(8)   Both the British and the French biologists had been searching Malaysia and
         Indonesia for the endangered monkeys.

(9)   Finally, the (British/French) spotted one of the monkeys in (Malaysia/Indonesia)
         and planted a radio tag on it.

                                                                        (Fraundorf et al., 2010)

In a later recognition memory task, participants remembered focus alternatives better if the focused element (here: *British/French* and *Malaysia/Indonesia*) had been produced with a contrastive accent ($L+H^*$) compared to a non-contrastive accent (H^*). However, memory for the focused element itself was not enhanced by a contrastive focus (Fraundorf et al., 2010). These findings indicate the importance of intonational focus for remembering the alternative set.

   In Spalek et al. (2014), the effect of different types of focus particles (inclusive *even* versus exclusive *only*) on the memory for focus alternatives was tested. A delayed recall task was performed by native speakers of German. Similar to Fraundorf et al. (2010), the auditory stimuli included a context passage (10) which was followed by a critical sentence (11).

(10)   Matthias receives a package with shirts, trousers, and jackets.
         He considered what he liked.   (context)

(11)   He kept (a) only/ (b) even/ (c) _ the shirts.   (critical sentence)

During the recall phase, participants were asked "What was in the package Matthias received?" Analyses revealed that the recall for focus alternatives (here: *trousers* and

jackets) was enhanced if the discourse had contained a focus particle. However, there was no significant difference in recall performance between the different particle types. Also, the focus particle did not improve memory for the focused element (here: *shirts*). The results indicate that a focus particle increases the salience of focus alternatives, making them easier to remember.

On the basis of this study, Koch and Spalek (in progress) created a similar experiment which used prosody instead of particles to mark focus. The auditory stimuli were constructed based on Spalek et al. (2014), see example (12) and (13).

(12)   Matthias receives a package with shirts, trousers, and jackets.
        He considered what he liked.   (context)

(13)   He kept the shirts/[SHIRTS]$_F$.   (critical sentence)

The results showed that the presence of intonational focus on the last element of the critical sentence (here: *shirts*) facilitates the recall of focus alternatives by native German speakers.

Furthermore, Koch and Spalek (in progress) found reduced overall recall and smaller focus effects on recall for male listeners compared to female listeners. The authors present evidence from the literature that women are more sensitive than men to emotional information conveyed through intonation (e.g., Hung & Cheng, 2014; Schirmer, Kotz, & Friederici, 2002; Wildgruber, Pihan, Ackermann, Erb, & Grodd, 2002) and they argue that women might generally outperform men in the exploitation of intonational cues to meaning. There is another explanation for the apparent gender difference: A number of recent studies (e.g., Bishop, 2016; Nieuwland, Ditman, & Kuperberg, 2010; Xiang, Grove, & Giannakidou, 2013) have shown that performance on the Autism-Spectrum Quotient questionnaire (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) is correlated with participants' performance in pragmatic language tasks. Even for a neurotypical population where individuals are far below a cut-off point for clinical diagnosis, autistic traits are variable and seem to play a role in communicative functioning (Bishop, 2016). Bishop's study is particularly relevant to our own research: He presented auditory sentences (in English) that were identical in all tested conditions, asking participants which sentence constituent they perceive as the most prominent one. Auditory stimuli were preceded by questions. If the question targeted the verb phrase, the verb was perceived as most prominent, whereas, if the question targeted the object, the object noun was perceived as most prominent. This is a kind of auditory illusion where context causes participants to observe differences in prominence that do not, in fact, exist in the stimulus. This effect was modulated by participants' scores on the Autism-Spectrum Quotient questionnaire such that participants with lower autistic traits were much more affected by context than participants with higher autistic traits. Autistic traits tend to be higher for men than for women (Baron-Cohen et al., 2001). Thus, an effect that looks like a gender effect, at first sight, might actually be an influence of autistic traits in disguise. Therefore, we decided to assess the participants' scores of the Autism-Spectrum Quotient questionnaire in our Vietnamese replication of the German study by Koch and Spalek (in progress).

2. Method

2.1. Participants

A total of 71 native Vietnamese participants were recruited via social network platforms. One participant was excluded from further data analysis because he misunderstood the task. The remaining participants were aged 18–39 years (M = 25.4, *SD* = 4.61).

Forty-six participants were female and 25 were male. Forty-two had a high school degree, 20 a bachelor's degree, and nine a master's degree. The educational level was not balanced across gender, but all of the participants completed a high school education in Vietnam. They were living in Germany for a minimum of one month and a maximum of 10 years (M = 2.74, *SD* = 1.89) at the time of the experiment. But all of them had lived in Vietnam until at least 15 years[2] old and spoke in the Northern dialect. Age of acquisition for German ranged between 13 years and 38.2 years (M = 21.87, *SD* = 5.30). Four participants had been speaking German for less than a year and two had not learned any German even though they were living in Germany. Participants were paid 12 Euro for their participation.

All speakers were bilinguals (even the two who did not speak any German, spoke English), and therefore, our results might not generalize to a monolingual population. For those participants who had been living in Germany for many years, it is even possible that language attrition might have taken place. In order to test for this, we looked at the influence of age of acquisition for German and length of residence in Germany on the results. Additionally, following the suggestion of an anonymous reviewer, we collected proficiency scores in a post-hoc survey (Section 2.2.2.2). We managed to collect scores for 59 (38 female and 21 male participants) out of the original 71 participants and will report whether language proficiency in German and Vietnamese affected the results. These proficiency scores revealed that even for the speakers with the highest proficiency in German, Vietnamese was still the stronger language and no attrition had taken place.

2.2. Materials

2.2.1. Main experiment

2.2.1.1. Auditory stimuli

Eighty stimuli (48 experimental items and 32 filler items) were created on the basis of the German stimuli used by Koch and Spalek (in progress). The stimuli in Koch and Spalek (in progress) were short stories consisting of three sentences. The first sentence introduced a protagonist and three list items from one taxonomic category (e.g., 'furniture,' 'tools,' 'fruits') followed by a prepositional phrase. For example, *Tamara had pearls, rubies, and sapphires in her vault.* The following sentence was a general statement about the protagonist (*She needed some money.*). In the last sentence, one of the three list items (*pearls, rubies, sapphires*) was repeated: *She sold the pearls.* This item is the *focused element.*

We used the same structure for the Vietnamese stimuli; see example (14). The list items in the first sentence were controlled for tone and number of morphemes: Twenty-four target stimuli and 16 filler stimuli included three list items with high tones (*ngang, sắc, ngã*) and 24 target stimuli and 16 filler stimuli included three list items with low tones (*huyền, hỏi, nặng*). Furthermore, for half of the target (24) and half of the filler stimuli (16), all three list items consisted of words with two morphemes (e.g., *nước mía* 'juice,' *sinh tố* 'smoothie,' *bia hơi* 'beer'), and for the other half, of words with one morpheme (e.g., *bí* 'pumpkin,' *khoai* 'potato,' *mướp* 'luffas'). The distribution of items with the same tone and morpheme number is given in **Table 1**.

The constraints were chosen to enhance comparability between the items and reduce interfering influences on memory performance such as word length. If possible, the German stimuli were translated into Vietnamese. In some cases, we had to create a more familiar

---

[2] Length of residence in Vietnam was calculated by subtracting the number of years participants had lived in Germany from their age. However, each participant was questioned whether he/she had lived in Vietnam until the age of 16, and all responded "yes." We assume that the discrepancy is caused by rounding.

Table 1: Distribution of tone and morpheme number of the items in the target and filler stimuli.

		Target	Filler
low tone	**monomorphemic**	13	6
	bimorphemic	11	9
high tone	**monomorphemic**	11	6
	bimorphemic	13	11

context for Vietnamese native speakers or had to change the list items because they did not conform to the tone and morpheme constraint. The structure of these translated stimuli might not represent natural storytelling. However, this approach was chosen because it reduces inter-item variability and makes the items comparable in an experimental setting. Whether the focused element in the last sentence was the first, second, or third from the list in the first sentence, was counterbalanced across items. Thus, we ensured that memory performance for the items was not an order effect.

(14)   a.   My   thấy có **khoai**, **bí**   và **mướp** ở khu rau
           My(name) see have potato, pumpkin and luffa   in area vegetable
           củ   trong siêu.thị.
           bulb in   supermarket
           "My looks at potatoes, pumpkins, and luffas in the vegetable section in the supermarket."   (context)

       b.   Cô.ấy nghĩ xem hôm.nay mình   định   nấu món gì.
           she   think see   today   oneself intend cook meal what
           "She considered what she wanted to cook today."   (context)

       c.   Cô.ấy đã   mua bí/[BÍ]$_F$.
           she   PST buy   pumpkin
           "She bought pumpkins."   (critical sentence)

The list items in each stimulus were not controlled for lexical frequency. In the German study (Koch & Spalek, in progress), frequency information was added to the analysis, and no effect was shown. Based on this finding, we assume that the influence of frequency is negligible in our task.

The stories were recorded by a native, female Vietnamese speaker of the Northern dialect. The speaker was unaware of the purpose of the experiment. The context sentences were recorded in two different versions: a narrow focus condition and a wide focus condition. The speaker was instructed to imagine that she was having a conversation with another person. She recorded two sentences (14-a) and (14-b) as if she were telling a story. Then, she had to imagine being asked, "What happened?" to record (14-c) like "She bought pumpkins." (i.e., the wide focus condition) as a response. Next, the speaker had to imagine that the other person did not properly understand and therefore, asked for confirmation: "Did she buy luffas?" The speaker then recorded (14-c) like "She bought PUMPKINS." as a correction (i.e., the narrow focus condition). Recording took place in a silent (but not sound-attenuated) booth. The digital auditory recorder (i.e., Philips Voice Tracer DVT2710 DNS Diktiergerät) was fixed at a distance of about 10cm from the speaker's mouth.

Based on a suggestion by an anonymous reviewer, the auditory stimuli (only targets, not the filler items) were submitted to a post-hoc naturalness rating. Four native speakers of the Northern dialect (female, age M = 34.5), who are living in Vietnam and have never been

to Germany, participated in the online rating. They rated the recorded stories on a 5-point scale (from 1 "not natural speech at all" to 5 "very natural speech"). Each participant heard 24 stimuli in the narrow focus and 24 stimuli in the wide focus condition. The stimuli were presented in a pseudo-randomized order such that no more than 3 stimuli of the same condition (narrow/wide focus) succeeded one another. The first two participants rated the target stimuli 1–24 in the narrow focus condition and 25–48 in the wide focus condition. For the third and fourth participant, we changed the conditions so that they rated the target stimuli 25–48 in the narrow focus condition and 1–24 in the wide focus condition. Thus, each stimulus received two ratings for each condition: The target stimuli have a mean value of 3.69 (*SD* = 0.69) in the narrow focus condition and of 3.75 (*SD* = 0.79) in the wide focus condition. Although the stimuli are not perceived as very natural, naturalness between the conditions does not differ.

In addition, we analyzed the auditory stimuli of the target items with Praat (Boersma & Weenink, 2009). Based on the studies which investigated focus intonation in Vietnamese (Jannedy, 2007; Michaud, 2004; Miller et al., 2015), we concentrated on the acoustic parameters duration, maximum and minimum pitch, and intensity. **Table 2** shows the means and standard errors for the parameters. The acoustic analysis was based on the

Table 2: Acoustic parameters of the focused elements in the critical sentences (e.g., *bí* 'pumpkins' in (14), target items only). The table shows the mean values (M) and standard errors (SE) for the narrow and wide focus condition as well as the *p*-value of the Welch two-sample t-test between the conditions.

Analysis interval **Condition**	Narrow focus		Wide focus		
	M	**SE**	**M**	**SE**	**p**
Focused element with tone *sắc* (n = 14)					
Duration	349.3	107.9	271.5	88.3	<.05*
Maximum pitch (Hz)	338.7	49.4	252.5	87.2	<.01**
Minimum pitch (Hz)	195.4	50.2	71.8	48.1	.2
Intensity (dB)	59.8	3.3	55.3	3.8	<.01**
Focused element with tone *huyền* (n = 13)					
Duration	387.4	98	340.8	140.4	.3
Maximum pitch (Hz)	202.8	27.7	182	48.1	.1
Minimum pitch (Hz)	159.5	42.9	149.4	43.5	.5
Intensity (dB)	59.1	3.8	54.3	4.2	<.01**
Focused element with tone *nặng* (n = 10)					
Duration	280	65.5	265.2	54.4	.5
Maximum pitch (Hz)	200.6	26.1	186.5	27.6	.2
Minimum pitch (Hz)	174.3	34	150.1	44	.1
Intensity (dB)	58.4	3.5	54.9	4	<.05*
Focused element with tone *ngang* (n = 9)					
Duration	337.1	90.4	305.9	47	.3
Maximum pitch (Hz)	239.6	22.3	201.4	45.7	<.05*
Minimum pitch (Hz)	206.1	36.3	177.3	27	<.1
Intensity (dB)	58.5	2.9	54.8	3.4	<.05*

syllable. Half of our items were monosyllabic words. For disyllabic words, the second syllable which received a different intonation in both focus conditions was used for the acoustic analysis. The target items consisted of 14 focused elements with tone *sắc*, 13 with tone *huyền*, 10 with tone *nặng*, 9 with tone *ngang*, 1 with tone *hỏi*, and 1 with tone *ngã*. Note that we grouped the tones into high (*sắc, ngang, ngã*) and low tones (*huyền, hỏi, nặng*). Therefore, the stimuli are not balanced for each of the Vietnamese tones and **Table 2** presents only the tones *sắc, huyền, nặng,* and *ngang*. For the tones *hỏi* and *ngã*, we had no basis for a comparison because there was only one stimulus for each tone.

Table 2 shows the mean values and standard errors for the acoustic parameters duration, maximum and minimum pitch, and intensity of the focused elements in the narrow and wide focus condition. The focused elements were analyzed separately for each tone (*sắc, huyền, nặng,* and *ngang*). The analysis with a Welch two-sample t-test showed that the focused elements in the narrow focus condition had significantly higher intensity values across all tones. This result supports the findings by Jannedy (2007) that focus is produced with higher intensity in Vietnamese. In addition, the high tones *sắc* and *ngang* had a significantly higher maximum pitch value in the narrow focus condition compared with the wide focus condition. Michaud (2004) also found an increase of F0 maxima in the production of focus. Furthermore, the duration of the focused element in the narrow focus condition was significantly longer for the tone *sắc* which is supported by the results in Jannedy (2007) and Miller et al. (2015). Note that the speaker chose to produce the focused element with a breathy voice in some stimuli in the wide focus condition. This could also lead to lower F0 maxima and intensity.

The differences between the pitch contours in both focus conditions are illustrated in **Figure 1**. Each graph represents the mean F0 over the focused element with a certain tone: *sắc* (1a), *huyền* (1b), *nặng* (1c), and *ngang* (1d). The graphs show that the words have a higher F0 mean value for each interval in the narrow focus condition compared to the wide focus condition across all tones. However, as illustrated in **Table 2**, only the pitch values of the tones *sắc* and *ngang* differ significantly. The analysis indicates that the speaker alternated the intonation for the words with different tones to varying degrees.

2.2.1.2. Recall questions
The questions for the recall task were based on the content of the stories. The corresponding question for the story in example (14) was:

(15) My   nhìn.thấy những loại   rau   củ   nào   trên giá.xếp
      My(name) see   some   type vegetable bulb which on   shelf
      rau   củ   ở cửa.hàng?
      vegetable bulb in supermarket
      "What types of vegetables did My see on the vegetable shelves in the supermarket?"

Forty-eight questions asked about the three list items. These were our target questions. In addition, we used 32 filler questions to ask for more general information about the story. These questions asked whether the protagonist of a certain story was female or male (16), what the action/situation was (17), or they were yes-no-questions (18):

(16) Một người   phụ.nữ hay một người   đàn.ông đã ở trong quán
      One human.CLF woman or   one human.CLF man   PST in inside shop
      cà.phê?
      coffee
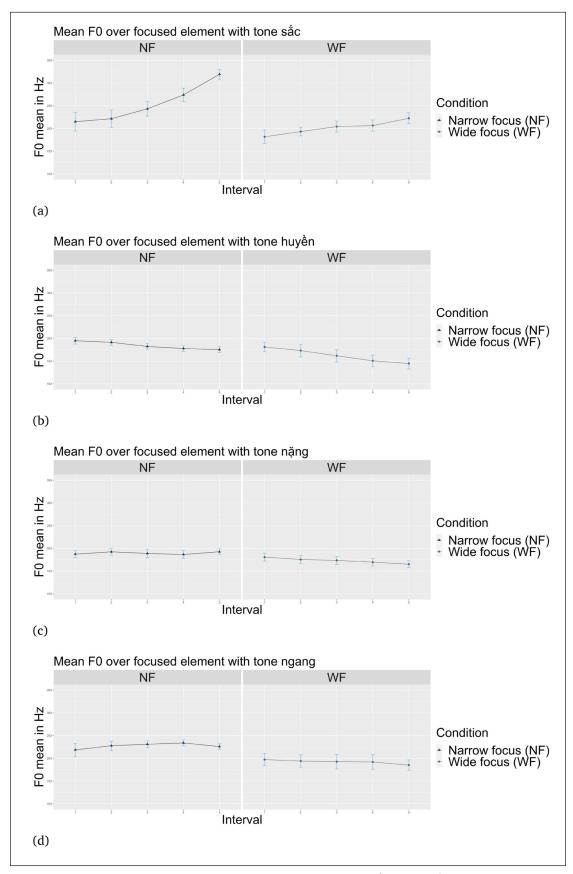   "Was a woman or a man in the coffee shop?"

Figure 1: Mean pitch contour of focused elements with tone *sắc* (a), *huyền* (b), *nặng* (c), *ngang* (d) in the critical sentences for both focus conditions. The left box demonstrates the contours in the narrow focus condition (triangles) and the right box illustrates the contours in the wide focus condition (dots). The blue whiskers indicate the standard error for each interval.

(17) Hoà   có   kế.hoạch gì   cho cả gia.đình vào mùa.hè?
      Hoà(name) have plan   what for all family in summer
      "What did Hoà plan for the whole family in the summer?"

(18) Có   phải Hân   tránh khu lộn   ở khu nhào.lộn?
      have right Hân(name) avoid area acrobatics in area playground
      "Did Han avoid the acrobatics area at the playground?"

The filler questions functioned as distractors and assessed whether the participants listened to the entire story.

Unlike Koch and Spalek (in progress), we did not ask for the names of certain protagonists in the stories. This decision was made after we received the feedback from the German participants that remembering names was the most difficult task in the experiment.

2.2.2. Predictor tasks

2.2.2.1. Autism-Spectrum Quotient questionnaire

We used a short version (AQ-k: see Freitag et al., 2007) of the Autism-Spectrum Quotient (AQ) questionnaire which was created by Baron-Cohen et al. (2001). The short version included 33 items instead of the original 50 items. The short version existed in German and was further translated into English. The English version was the basis for our translation into Vietnamese (see supplementary material of this article).

2.2.2.2. Language experience and proficiency questionnaire

The Language Experience and Proficiency Questionnaire (LEAP-Q) designed by Marian, Blumenfeld, and Kaushanskaya (2007) is a reliable and efficient questionnaire of bilingual language profiles. It is applied to collect self-reported language proficiency data from speakers aged 14 to 80 and contributes to predicting a relationship between speakers' bilingual language status and their measured language behaviors. We used the Vietnamese version translated by Phạm and Nguyễn (see Marian et al., 2007). The questionnaire was administered post-hoc: The second author of this article contacted all participants again and asked them to turn in a completed version of the questionnaire online.

2.3. Design

We created two supplementary experimental lists. The 80 auditory stimuli (48 target and 32 filler items) were distributed evenly across the two experimental conditions (narrow versus wide focus). Items that were assigned wide focus in the first list were assigned narrow focus in the second list and vice versa. Half of the participants were presented with List 1 and half with List 2. An equal number of males and females were tested on each list. A list was divided into eight blocks. In each block, ten items were presented: six target items and four filler items. While creating the lists, we made sure that the blocks were balanced for focus condition (five items with narrow focus and five items with wide focus). For the other structural features, we spread the items as evenly as possible across type (target/filler), tone (high/low), and morpheme number (1/2). No more than three items of the same structure occurred in a row. A given block did not always contain equal numbers of words with 1-morpheme and 2-morphemes or of words with high and low tones. The relation between morpheme number (1-morpheme versus 2-morphemes words) varied from 4:6 to 6:4. For the tones, the relation of words with high versus low tones diverged between 3:7 to 7:3. In addition, we took care that each type of response to the filler questions (yes/no, gender of the protagonist, general information) occurred at least once and no more than twice in one experimental block. The order of the items was the same for each participant. This approach followed the procedure in Koch and Spalek

(in progress). However, we ensured that stories with similar topics such as 'zoo animals' and 'farm animals' did not occur in the same experimental block.

Table 3 illustrates the stimulus distribution in the first experimental block.

2.4. Procedure

Participants signed an informed consent form and an information form about data protection. After the forms were filled out, the experimenter (the first author of this article who is a German native speaker) gave the participant an oral instruction of the task. The communication between the experimenter and participant was done in German or English, depending on the participant's preference. In addition, the same instruction was presented in written Vietnamese on the computer screen. The instruction included a short description of the task and the structure of the experiment. It explained that participants had to listen to stories via headphones. They were told to listen carefully because they had to answer questions about the stories after hearing several of them. An example was also given to prevent misunderstandings. The instructions made clear that their answers were recorded and that they had 22 seconds to answer each question. They were instructed to indicate if they did not know an answer by saying *Tôi không biết.* "I don't know." Participants did not receive feedback on whether or not their answer was correct. After the instructions, a practice block with six practice items followed. If the participants had any questions about the experiment procedure, they could ask them after the practice block. The interaction between the experimenter and the participants was short and was intended to make sure that the participant felt comfortable in the formal setting of the laboratory. The participant interacted mainly with the computer where the language mode was completely in Vietnamese.

One experimental block was structured as follows: In the first phase (encoding), the participants heard 10 stories through a Sennheiser PC8 headset with an integrated microphone. A fixation cross in the center of the screen appeared while a story was played. The second phase was the test phase (recall phase) in which the participants had to answer the questions. One question per story was presented on the monitor (10 questions per block). The questions were presented in the same order as the stories. Following the recall phase, the participants had to perform a ten-step n-backward counting task

Table 3: Distribution of stimuli in one experimental block based on their structure and content.

ID	Block	Stimuli type	Tone	Number of morphemes	Focus condition list 1	Focus condition list 2	Topic	Focused element	Question answer type
F24	1	filler	high	2	WF	NF	school	1	yes/no
F13	1	filler	low	2	NF	WF	hobbies	3	yes/no
T18	1	target	high	1	NF	WF	vegetables	2	list items
T25	1	target	low	1	WF	NF	wood	3	list items
F68	1	filler	high	2	NF	WF	mushrooms	3	general
T09	1	target	low	1	WF	NF	sports	1	list items
T03	1	target	high	2	NF	WF	beverages	1	list items
F53	1	filler	high	1	WF	NF	choirvoices	1	male/female
T41	1	target	low	2	WF	NF	professions	3	list items
T01	1	target	high	1	NF	WF	clothes	2	list items

which was presented in Vietnamese on the computer screen. They were instructed to count backward, for example, from 30 in steps of 3. Participants intuitively counted in their native language which suggests that they stayed in this language mode throughout the experiment. Although they had to speak the numbers out loud, the counting was not recorded and controlled for accuracy. However, the experimenter ensured that they executed this task because it should reduce interference effects between recurring taxonomic categories. **Table 4** presents a schematic overview of the experimental procedure.

Stimulus presentation and recording of the answers was controlled by Neurobehavioral Systems Presentation software (Version 16.5). The stories in the encoding phase were played with a silent interval of 4 seconds between them. After the presentation of 10 stories, the recall phase was initiated automatically. A short announcement indicated that the question phase was soon to start (display duration 2 seconds). The question followed after a central fixation cross was presented for 500 milliseconds on the screen. After the question had disappeared (display duration 3 seconds), a '#'-symbol was presented to indicate that the participants could respond to the question now. They were instructed to answer aloud as soon as possible when the '#'-symbol appeared. However, the actual recording started from the onset of the question. If a participant answered while the question was still displayed on the screen, we did not lose any information. When participants had finished responding before the next question appeared, they could jump to the following question by pressing the space key.

Following the main experiment, participants did a short sentence reading task that will be reported elsewhere. After the sentence reading task, participants filled out the Autism-Spectrum Quotient questionnaire.

Finally, the experimenter asked the participants for basic demographic information and whether they employed any strategies in the main experiment. The participants were informed about the aim of the study after all tasks were completed. An entire testing session lasted about 70 minutes.

2.5. Data annotation

Recall accuracy was coded for the recorded answers. The answers were independently annotated by the first and second author of this paper. The two annotators agreed in 93.9% of the cases ($\kappa = 0.877$). For the target items, each word of the set (focused element, first focus alternative, second focus alternative) was coded binary: Either it was 'recalled' (coding: '1') or 'not recalled' (coding: '0'). That is, a given item (story) yielded two responses for the alternatives and one response for the focused element. We used a strict and lenient coding. In the strict coding, a '1' was only assigned if the word was recalled in its original form. If a participant used a synonym, see example (19), or did not name the classifier, as in example (20), the words were annotated as '0' in the strict, but as '1' in lenient coding. Although participants were not explicitly instructed to recall the list items with classifiers, they commonly remembered the items with classifiers. They also rarely thought of synonyms. In natural speech, the classifiers are sometimes omitted if the context is clear. Thus, the meaning of the words does not change; see example (20).

(19)   a.   máy.xúc   (strict coding)
              shovel/dig
              'excavator'

         b.   máy.đào   (lenient coding)
              dig
              'excavator'

Table 4: Structure of an exemplary experiment block with English translation. A complete list of the stimuli which were used in the study can be found in the supplementary material of this paper.

Stimuli	Trial Type	Vietnamese transcript	English transcript
Encoding Phase of Auditory Stimuli			
Item 1	Filler	Tuệ giải thích về các chất bao gồm các-bon, ba-zơ và a-xít trong một giờ hoá học. Cô ấy thực hiện một bài thuyết minh về chúng. Cô ấy đã giải thích rõ về các-bon.	Tue explained the substances including carbon, basalt, and acid in a chemistry lesson. She made a presentation about them. She clearly explained everything about carbon.
Item 2	Filler	Xuyên đọc quảng cáo về các khoá học gồm dịch thuật, làm vườn và hội hoạ ở một trung tâm học nghề. Anh ấy muốn mở rộng và nâng cao các kỹ năng của bản thân. Anh ấy đã tham dự khoá học hội hoạ.	Xuyen read leaflets about courses including translation, gardening, and painting at an apprenticeship center. He wanted to expand and improve his skills. He attended the course on painting.
Item 3	Target	My thấy có khoai, bí và mướp ở khu rau củ trong siêu thị. Cô ấy nghĩ xem hôm nay mình định nấu món gì. Cô ấy đã mua bí.	My looks at potatoes, pumpkins, and luffas in the vegetable section in the supermarket. She considered what she still had at home. She bought pumpkins.
Item 4	Target	Lâm làm một nghiên cứu nhỏ về chồn, lừa và ngựa khi học môn sinh học. Cô ấy sưu tầm được rất nhiều tài liệu. Cô ấy đã viết xong về ngựa.	Lam did a little research on minks, donkeys, and horses when she studied biology. She collected a lot of documents. She wrote about horses.
Item 5	Filler	Việt trồng nấm hương, nấm rơm và nấm mối. Anh ấy biết các loại nấm ấy rất tốt cho sức khoẻ. Anh ấy đã bán được rất nhiều nấm mối.	Viet grew lentinula edodes, volvariella volvaceas, and macrolepiota albuminosas. He knew that they were good for health. He sold a lot of macrolepiota albuminosas.
Item 6	Target	Nam thử các bài tập với vòng, tạ và xà trong phòng tập thể dục. Anh ấy muốn chọn xem bài tập nào anh ấy muốn thực hành. Anh ấy đã rất thích bài tập với vòng.	Nam discovered hoops, dumbbells, and beams in the gym. He contemplated on which exercises he would like to do. He picked out the hoops.
Item 7	Target	Lan mua nước mía, sinh tố và bia hơi ở một cửa hàng đồ uống. Cô ấy muốn làm dịu cơn khát của mình. Cô ấy đã uống nước mía.	Lan bought sugarcane juice, smoothies, and beer at the beverage shop. She wanted to quench her thirst. She cooled sugarcane juice.
Item 8	Filler	Nghị kiểm tra các nhóm thanh nhạc cho nam bao gồm thấp, trung và cao. Ông ấy từng là giảng viên thanh nhạc. Ông ấy đã đánh giá tốt giọng nam thấp.	Nghi examined vocal groups for men including low, middle, and high vocals. He was a vocal coach. He appreciated the low vocals.

Stimuli	Trial Type	Vietnamese transcript	English transcript
Item 9	Target	Thịnh gặp nhiều thợ tiện, thợ nguội và thợ dệt ở hội chợ nghề nghiệp. Cô ấy muốn có một gian hàng riêng của mình. Cô ấy đã nói chuyện với thợ dệt.	Thinh met the bricklayer, the electrician, and the weaver at the construction site. He wanted to examine the work. He talked to the weaver.
Item 10	Target	Thanh nhận được một gói đồ, trong đó có áo, khăn và tất. Anh ấy thích màu xanh lá cây nhất. Anh ấy đã thử quàng khăn.	Thanh received a package with shirts, scarves, and socks. He liked the green color very much. He tried on scarves.
Recall Phase with Orthographical Stimuli			
Question 1	Filler	Tuệ đã làm gì trong giờ hoá học?	What did Tue do in the chemistry lesson?
Question 2	Filler	Xuyên đã tham gia lớp học hội hoạ phải không?	Did Xuyen participate in the painting class?
Question 3	Target	My nhìn thấy những loại rau củ nào trên giá xếp rau củ ở cửa hàng?	What types of vegetables did My see on the vegetable shelves in the supermarket?
Question 4	Target	Lâm tìm kiếm những loại động vật nào trong rừng?	What kind of animals was Lam looking for in the forest?
Question 5	Filler	Việt đã bán được nhiều nấm rơm có phải không?	Did Viet sell many mushrooms?
Question 6	Target	Nam đã khám phá ra những loại dụng cụ nào trong phòng thiết bị tập thể dục?	What kind of equipment did Nam discover in the gym room?
Question 7	Target	Lan đã mua loại nước giải khát nào ở siêu thị?	Which beverage did Lan buy at the supermarket?
Question 8	Filler	Một người đàn ông hay một người phụ nữ đã khen ngợi giọng nam thấp?	Did a man or a woman appreciate the low vocals?
Question 9	Target	Thịnh gặp những ai ở hội chợ nghề nghiệp?	Who did Thinh meet at the construction site?
Question 10	Target	Thanh nhận được những gì trong gói đồ?	What did Thanh get in the package?
10-step backward counting task			
8-decrement		80–72–64–56–48–40–32–24–16–8–0	

(20) a. thợ.hồ   (strict coding)
         mason
         'mason'

    b. hồ   (lenient coding)
         mason
         'mason'

The agreement between the two annotators was calculated based on the strict coding. If the two annotations mismatched, the first author listened to the participant's answer again to decide on the right coding. Synonyms were mainly coded by the second author

since she is a native Vietnamese speaker. If a participant indicated that he/she did not know the answer to the recall question, all three words of that trial were coded as 'not recalled' ('0').

The filler items were annotated by the second author because the participants' answers were more complex in some cases. The answers were also coded binary as 'right' ('1') and 'wrong' ('0'). However, the participants had a greater freedom in their answers.

3. Results

We excluded the data from three participants who recalled less than 20% of the words in the critical conditions. No participants were excluded due to their performance in the filler questions. **Table 5** presents the descriptive results. Koch and Spalek (in progress) observed a beneficial effect of focus intonation for the recall of alternatives, but not for the recall of the focused element. Therefore, we also present the descriptive data for alternatives and the focused element separately in **Table 5**. In addition, the data are split by participant gender which was also a factor in our statistical model (see below). Results are presented for strict coding. The results for lenient coding are qualitatively identical; they just comprise more recalled cases. Given that strict coding leaves less scope for interpretation on whether the participant had intended the correct referent or not, all analyses are based on strict coding.

Descriptively, we observe a small memory benefit for alternatives (for women: M = 46.9% contextual alternatives recalled in the narrow focus condition compared to M = 44.4% contextual alternatives recalled in the wide focus condition). As in Koch and Spalek (in progress), this benefit is more pronounced for the female sample. In fact, for the male sample, the descriptive effect reverses, with more alternatives recalled in the wide focus condition (39.0%) than in the narrow focus condition (37.1%). Unlike Koch and Spalek (in progress), the data pattern for recall of the focused element is identical to the pattern for alternative recall (for women: M = 67.2% focused elements recalled in the narrow focus condition compared to M = 64.0% focused elements recalled in the wide focus condition). Again, the pattern in the male sample was reversed (M = 59.6% narrow focus condition compared to M = 62.0% wide focus condition).

Statistical analyses were done using logistic mixed effects models in the statistical computing environment R with the lme4 package (Bates, Maechler, Bolker, & Walker, 2014). Given the parallel results for alternatives and the focused element, we did not start with separate analyses but included word type (alternatives versus focused element) as a fixed factor to the model to determine whether separate analyses for the focused element and alternatives were statistically justified. Other than that, our initial model was identical to the one used by Koch and Spalek (in progress), including fixed effects for condition (narrow versus wide focus), word type (alternative versus focused element), participant gender (male versus female), and their interactions. An additional fixed effect was used to

Table 5: Descriptive results (in percent) for alternative recall split by condition (narrow versus wide focus), type (alternatives versus focused element), and participant gender (men versus women versus all participants). Standard deviations are presented in brackets.

	Alternatives		Focused element	
	Narrow focus	**Wide focus**	**Narrow focus**	**Wide focus**
Men	37.1 (48.3)	39.0 (48.9)	59.6 (49.1)	62.0 (48.6)
Women	46.9 (49.9)	44.4 (49.7)	67.2 (46.9)	64.0 (48.0)
All Participants	43.5 (49.6)	42.6 (49.5)	64.6 (47.8)	63.3 (48.2)

model (centered) trial number, that is, the position of an item in the experiment. Condition, word type, and participant gender were sum-coded. Random intercepts were included for participants, items (i.e., a given story), and words (i.e., the word to be recalled). Random slopes were added for trial number on the participant intercept and for participant gender on the item intercept. We confirmed via likelihood ratio tests that the complex random effects structure from Koch and Spalek (in progress) indeed described the data best. Next, we tested, again with likelihood ratio tests, whether the interactions of word type and condition and word type and gender were really necessary. Neither of them improved model fit ($\chi^2(2) < 1, p = 0.92; \chi^2(1) = 1.92, p = .16$, respectively). By contrast, adding the main effect for word type did improve model fit significantly ($\chi^2(1) = 62.27, p < .001$). The model fit was additionally improved when we modeled trial number not as a linear predictor but as a polynomial one ($\chi^2(4) = 31.22, p < .001$).

Having thus identified the best model based on the experimental design, we tested whether the addition of predictor variables further improved model fit. Adding participants' scaled and centered autism scores did improve model fit significantly ($\chi^2(1) = 4.05, p = .04$). However, adding the interaction between condition and autism scores did not lead to any further improvement ($\chi^2(1) < 1, p = .59$). Finally, we tested whether the effect of condition interacted with participants' age of acquisition for German and/or their length of residence in Germany. Both variables were scaled and centered and neither the interaction of condition by age of acquisition ($\chi^2(2) = 1.17, p = .56$) nor the interaction of condition by length of residence ($\chi^2(2) = 2.75, p = .25$) improved model fit. **Table 6** shows the results of the final model.

Table 6: Fixed effect estimates and variance estimates for GLMER of recall for list items (recall~condition * gender + autism_score + WordType + poly(NumInExpC,2) + (1+gender|item) + (1|word) + (1+poly(NumInExpC,2)|participants), n = 9648, log-likelihood = −5322.4). Coding scheme for gender, condition, and word type = sum coding.

Random effects

Groups Name	Variance	SE	Corr	
word (Intercept)	0.4302	0.6559		
participants (Intercept)	0.4942	0.7030		
item (Intercept)	0.8851	0.9408		
gender	0.1605	0.4006	−0.22	
poly(NumInExpC, 2)1	757.6458	27.5254	0.10	
poly(NumInExpC, 2)2	492.6876	22.1966	−0.20	0.02

Number of obs: 9648, groups: word, 144; participants, 67; item, 48

Fixed effects

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	0.12718	0.17542	0.725	0.5
focus condition	0.01514	0.05135	0.295	0.8
gender	0.39349	0.19665	2.001	<.05*
autism_score	0.18910	0.08862	2.134	<.05*
word type	−1.19963	0.12839	−9.344	<.0001***
focus condition:gender	0.28637	0.10271	2.788	<.001**
poly(NumInExpC, 2)1	34.02253	9.20417	3.696	<.0001***
poly(NumInExpC, 2)2	−14.86748	12.15805	−1.223	0.2

To summarize the results: There was no main effect of focus condition, but a significant interaction of focus condition and participant gender. A main effect of participant gender reflected that men recalled fewer words than women overall. The effect of the autism scores was significant such that, for participants with higher autistic traits, recall was improved. The effect of word type showed that the focused element was recalled significantly more often than alternatives. Only the linear component of the predictor of trial number was significant, suggesting that participants became better during the course of the experiment. We resolved the significant interaction of condition by participants' gender by looking at the effect of condition separately for male and female participants. **Table 7** presents the results for the men; **Table 8** presents the results for the women. **Figure 2** gives an overview of the results for the entire sample and men and women separately.

The comparison between men and women shows that focus intonation improves recall for women but not for men, in line with the findings by Koch and Spalek (in progress). Men do show an effect of word type, that is, more recall for the focused element than for alternatives. The same is true for women. Despite our best efforts, we did not manage to test a balanced sample of men and women. Thus, the results of 44 women and 23 men were included in the final analysis. In order to test whether the difference between men and women could simply be due to differences in statistical power, we carried out an additional analysis with a smaller female sample, namely the first 23 women who were tested in this study. Even for this reduced female sample, the effect of focus condition was still significant (B = 0.21, |z| = 2.52, p = .01).

Because all three authors of the present paper are currently based in Berlin, we had tested native speakers of Vietnamese who were currently living in Berlin. Even though we invited only speakers who had lived in Vietnam until at least the age of 15 and for

Table 7: Fixed effect estimates and variance estimates for GLMER of recall for list items for men (recall ~ condition + autism_score + WordType + poly(NumInExpC, 2) + (1 | item) + (1 | word) + (1 + poly(NumInExpC, 2) | participants), n = 3312, log-likelihood = −1856.8). Coding scheme for condition and word type = sum coding.

Random effects

Groups Name	Variance	SE	Corr	
word (Intercept)	0.3912	0.6255		
item (Intercept)	0.9886	0.9943		
participants (Intercept)	0.3252	0.5702		
poly(NumInExpC, 2)1	230.6333	15.1866	0.50	
poly(NumInExpC, 2)2	74.6809	8.6418	0.41	−0.16

Number of obs: 3312, groups: word, 144; item, 48; participants, 23

Fixed effects

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	−0.02837	0.19977	−0.142	0.9
focus condition	−0.12462	0.08319	−1.498	0.1
autism_score	0.09874	0.14454	0.683	0.5
word type	−1.28709	0.14476	−8.891	<.001***
poly(NumInExpC, 2)1	25.33997	9.76681	2.594	<.01**
poly(NumInExpC, 2)2	−9.75502	9.09338	−1.073	0.3

Table 8: Fixed effect estimates and variance estimates for GLMER of recall for list items for women (recall~ condition + autism_score + WordType + poly(NumInExpC,2) + (1|item) + (1|word) + (1+poly(NumInExpC,2)|participants), n = 6336, log-likelihood = −3534.6). Coding scheme for condition and word type = sum coding.
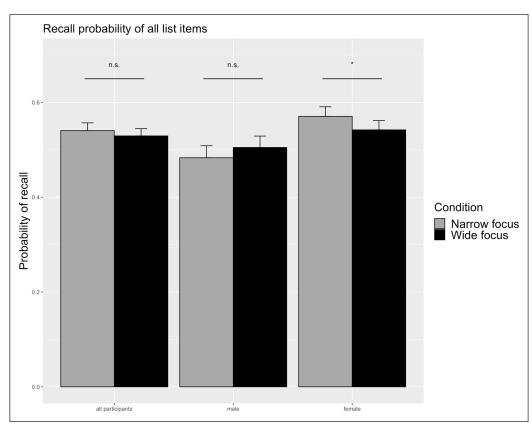
Random effects

Groups Name	Variance	SE	Corr	
word (Intercept)	0.3718	0.6097		
item (Intercept)	0.8487	0.9212		
participants (Intercept)	0.5776	0.7600		
poly(NumInExpC, 2)1	507.4210	22.5260	−0.03	
poly(NumInExpC, 2)2	412.3018	20.3052	−0.31	0.11

Number of obs: 6336, groups: word, 144; item, 48; participants, 44

Fixed effects

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	0.30509	0.18658	1.635	0.1
focus condition	0.15700	0.05994	2.619	0.01**
autism_score	0.19222	0.10657	1.804	0.07 .
word type	−1.13655	0.12708	−8.943	<.001***
poly(NumInExpC, 2)1	25.06059	10.09887	2.482	<.05*
poly(NumInExpC, 2)2	−11.13096	10.88942	−1.022	0.3

Figure 2: Results of the recall probability of all list items for all, male, and female participants in both conditions (narrow versus wide focus). The error bars indicate the standard errors. The asterisks above the bars illustrate the significant effects for each sample (* = p < .05; n.s. = not significant).

whom Vietnamese was their strongest language, our participant sample is still special in two ways that are relevant for the present study: First, they might have learned the function of focus accents in German and unconsciously transferred this to their processing of Vietnamese, that is, showing transfer from their second language to the first language. Second, they might have been living in Germany and using the German language for so long that language attrition for Vietnamese might have set in. The influence of L2 on L1 would be reflected in a stronger influence of intonation focus on recall performance for those participants with the highest German proficiency, that is, an interaction of German proficiency with focus condition. L1 attrition might show up in generally poorer performance for people with the lowest Vietnamese proficiency, that is, in a main effect of Vietnamese proficiency on recall performance.

We had obtained proficiency scores (Section 2.2.2.2) for 59[3] out of the original 71 participants (58 of these participants were part of the final data analysis). First, we ran the statistical model on this reduced data set and still observed the relevant effects: an interaction of condition with participant gender ($B = 0.25$, $|z| = 2.22$, $p = .03$), and main effect of gender ($B = 0.50$, $|z| = 2.48$, $p = .01$), word type ($B = -1.23$, $|z| = 9.67$, $p < .001$), and autism score ($B = 0.25$, $|z| = 2.64$, $p < .01$). The linear component of the predictor for trial number was also significant ($B = 31.52$, $|z| = 2.75$, $p < .01$). We tested for an influence of German proficiency by adding an interaction term for German proficiency and focus condition. This did not improve model fit ($\chi^2(2) < 1$, $p = .84$). We tested for an effect of Vietnamese attrition by adding Vietnamese proficiency to the model, which, again, did not improve model fit ($\chi^2(1) = 2.03$, $p = .15$). Thus, while it would certainly be important to replicate these findings with a truly monolingual sample tested in Vietnam, the data suggest that our results can neither be explained by L2 to L1 transfer nor by L1 attrition.

4. Discussion

In this paper, we aimed to test whether a contrastive intonation contour facilitates the recall of focus alternatives in a tonal language, namely Northern Vietnamese. Our experiment was based on the study by Koch and Spalek (in progress) with German participants. After listening to auditory discourses, participants were prompted to recall the list items in our stimulus set. The last element in the story was either realized with a contrastive focus intonation or without. The focus manipulation in the auditory sentences was achieved by presenting the person recording these sentences with contextual prompts, leading to either narrow or wide focus realization. As it turned out, the exact prosodic realization of narrow focus depended on the lexical tone of the critical word.

Descriptively, participants remembered both the focus alternatives and the focused element better in the narrow focus condition than in the wide focus condition. In addition, we found a significant interaction of participant gender and the focus condition: Women recalled more items when the focused element had been presented with a contrastive intonation contour whereas men did not show this faciliatory effect. In fact, descriptively, men even showed the opposite effect. We therefore replicated the results presented in Koch and Spalek (in progress). In contrast to previous research (Koch & Spalek, in progress; Spalek et al., 2014), we observed that contrastive intonation not only improved recall of contextual alternatives but also recall of the focused element itself.

---

[3] 38 female and 21 male participants took the LEAP-Q.

In the following, we discuss our findings within the context of focus alternative recall as well as gender differences in how focus is processed. Furthermore, we examine our acoustic analysis of the focused element in our target sentences with regard to the question of whether intonational focus exists in Vietnamese.

4.1. Recall of focus alternatives and focused element

The findings of our study indicate that focus alternatives in Vietnamese texts are remembered better if a contrastive focus intonation is present. This supports the results of studies with non-tonal languages such as English (Fraundorf et al., 2010) or German (Koch & Spalek, in progress). Fraundorf et al. (2010) showed that the recognition of focus alternatives was facilitated when the focused element was produced with an $L + H^*$ intonation contour. Although Fraundorf et al. (2010) had only one focus alternative per item, we could replicate their findings with more list items and a task that is arguably more difficult, namely recall. The effect was still apparent in our data which provides further evidence that focus marking makes alternatives more salient. In addition, our results support the assumption that focus calls listeners' attention to members of an alternative set (Rooth, 1992).

However, in comparison to the effect of conventionalized sensitivity to focus (see Beaver & Clark, 2009) for particles such as *only* and *even*, the influence of intonational focus seems to be less powerful and subject to stronger individual variation. In the German study (Koch & Spalek, in progress), on which the present experiment is based, the authors found that participants' gender had an impact on the processing of focus in that men showed a much smaller and non-significant memory benefit. Therefore, for the present study, we had tried to test a gender-balanced sample to further investigate those differences. Koch and Spalek (in progress) observed a benefit of 3.4% in the narrow focus condition. In contrast, the study with Vietnamese participants showed a benefit for remembering alternatives of 0.9% in the same condition. The comparison of women's memory in both studies revealed an even smaller benefit in the Vietnamese study (2.5%) than the 5% in Koch and Spalek (in progress). Although the distribution of female and male participants is not even (65.7% female) in our study, we replicated the interaction of focus by gender and showed that a memory benefit was only present in the female group. This memory benefit for women was still present when the female sample was reduced in size to match the male sample. We assume that our results illustrate a general tendency that supports the findings in Koch and Spalek (in progress) and will further discuss this finding in the next section.

Interestingly, our analysis showed a parallel pattern between the focused element and focus alternatives which was absent in other studies (Fraundorf et al., 2010; Koch & Spalek, in progress). In Fraundorf et al. (2010), the recognition of the focus alternative was enhanced when the focused element was presented with a contrastive intonation. However, the recognition of the focused element did not show this faciliatory effect. In a cross-modal lexical decision task, Yan and Calhoun (2019) investigated whether priming effects for focused words and their alternatives were increased with prosodic or syntactic focus marking in Mandarin Chinese. The study used a set of sentence stimuli that were presented auditorily. The grammatical subjects of the sentences were the primes. Priming effects for the subject as well as its contextual alternatives were increased by the presence of prosodic focus on the subject. In contrast, syntactic focus marking did not increase the priming effect for the focused element. The findings in Yan and Calhoun (2019) that both contextual alternatives and the focused word itself are primed by prosodic focus are in line with our results. It might be the case that Mandarin Chinese and Vietnamese behave differently compared to Indo-European languages in this regard. Further research needs

to test why prosodic focus improves the recall of the focused element and whether this might constitute a genuine difference between the processing of intonation focus in tone languages and languages without lexical tones.

4.2. Gender differences in focus processing

Our study replicated the finding from Koch and Spalek (in progress) that intonational focus marking influences the recall of alternatives only in female participants.

One possible explanation might be that our stimuli were read by a female speaker and that male listeners were therefore less sensitive to the prosodic information contained in it. However, while it is known that men and women differ in their realization of prosody (e.g., Anderson, Hiramoto, & Wong, 2007)[4], this explanation implies that men are less able to process prosodic cues provided by women. This is a question that needs to be tested empirically by replicating this study with a male speaker. However, given that Anderson et al. (2007) also showed that men were able to imitate women's use of prosody and vice versa, it seems improbable that men are not aware of focus marking cues used by female speakers. In fact, if this turned out to be true, after all, it would have rather far-reaching consequences for the study of language comprehension.

Another explanation for the gender effect is discussed in Koch and Spalek (in progress): Women and men might differ in how they process intonational cues. Some studies suggest that women could have an advantage in processing pragmatic information conveyed by prosody. Gender differences in the interpretation of intonation have been examined most frequently in investigations on emotional speech processing (e.g., Hung & Cheng, 2014; Schirmer et al., 2002; Schirmer, Striano, & Friederici, 2005; Schirmer, Zysset, Kotz, & von Cramon, 2004; Wildgruber et al., 2002). More closely related to our experiment are findings of gender differences obtained from studies on the interpretation or production of coherent narratives (e.g., Frank, Baron-Cohen, & Ganzel, 2015; Kaiser, Kuenzli, Zappatore, & Nitsch, 2007; Kansaku, Yamaura, & Kitazawa, 2000). In our experiment, participants listened to ten stories with various topics and had to answer questions afterwards. Female participants might have recognized and processed the intonational cues differently than male participants. In the introduction (Section 1), we had also speculated that the observed gender effect might be an effect of autistic traits in disguise. However, the data do not support this: Adding participants' scores from the Autism-Spectrum Quotient questionnaire (Baron-Cohen et al., 2001) did improve model fit, but only as a main effect. It did not interact with focus condition, that is, autism scores did not affect the size of the influence of prosodic information on recall.

Furthermore, it has been shown several times that women have an advantage in episodic memory tasks with verbal stimuli (e.g., Herlitz & Rehnman, 2008). In contrast, men seem to have a better performance than women in tasks that focus on episodic memory with visuospatial materials (Herlitz & Rehnman, 2008). Therefore, memory operations might be different in female and male participants which could lead to performance variation. The results of our study are in line with the finding that women outperform men in memory for verbal material. Thus, the memory benefit for the female listeners might stem from a combination of increased sensitivity to prosodic cues coupled with improved memory for verbal stimuli. On the one hand, the gender difference speaks against a grammaticalization of prosodic focus since we would expect grammaticalized functions to hold across the entire population. On the other hand, it might be that women drive the grammaticalization process.

---

[4] We thank an anonymous reviewer for pointing out this study to us.

4.3. Focus production in Northern Vietnamese

The results of our study demonstrate that the processing of prosodic focus in North Vietnamese and non-tonal languages like German or English is similar. It seems to be the case that Vietnamese uses intonation to communicate pragmatic functions despite its complex tone system. This is in line with findings from Jannedy (2007) that a local change in F0 indicates different focus markings, for example, subject- and verb-focus.

The recording of a question-answer paradigm with two native Northern Vietnamese speakers in Jannedy (2007) illustrates an increase of duration, intensity, and F0 to mark focus. In our stimuli, which were recorded by a female Northern Vietnamese speaker, we found comparable prosodic realization for the focused elements. Our speaker used syllable lengthening as well as a higher pitch and intensity in the narrow focus condition compared to the wide focus condition throughout all focused elements. However, in contrast to Germanic languages, we cannot pinpoint a particular accent like the $L+H^*$ accent that marks focus. Instead, the particular strategy for marking intonation focus in Vietnamese seems to depend on the tone of the critical word. However, our data do show that listeners are sensitive to the prosodic differences between wide and narrow focus marking in a similar way as reported previously for German.

Although speakers might vary in how they produce focus (see Brunelle, 2017), there seems to be an overall tendency to indicate focus marking with intonation (Ip & Cutler, 2017). As of yet, it is not clear whether female and male speakers produce focus in the same way.

5. Conclusion

By replicating a study originally designed for native German speakers (Koch & Spalek, in progress), we shed light on the influence of prosodic focus intonation on language processing in Northern Vietnamese. Although Vietnamese has a complex tone inventory, native speakers seem to process intonational cues in a similar way to non-tonal language speakers. Our results support the assumption that tonal languages can use intonation to structure information. The significance of focus alternatives in a given context can be marked with a higher pitch, duration, and intensity on the focused element in Northern Vietnamese.

This contrasting intonation contour appears to be beneficial to women's memory performance, while male participants showed no memory improvement between the two conditions. This finding suggests that women and men might process prosodic cues differently.

Abbreviations

The following abbreviations were used in the text:

$L+H^*$: rising pitch accent from low (L) to high (H) tone
L: low (L) boundary tone

The following abbreviations were used in the examples:

SG: singular
F: focus
PST: past
CLF: classifier

Additional Files

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

Acknowledgements

Competing Interests

The authors have no competing interests to declare.

Author contributions

Annika Tjuka is a research associate in the FAHMRRR project and managed the study. She revised the Vietnamese stimuli and prepared them for the Presentation script which she adapted with support from Carsten Schliewe. In addition, Annika Tjuka performed the acoustic analysis and set up the online rating study for the Vietnamese stimuli. She also tested the 71 participants and evaluated the Autism-Spectrum Quotient questionnaire.

Huong Thi Thu Nguyen, a PhD candidate at Humboldt University of Berlin, is supported by the Ministry of Education and Training, Vietnam. She translated our material into Vietnamese and advised the project as a native Vietnamese speaker. Furthermore, she annotated the recorded data of the main experiment in parallel with Annika Tjuka. Huong Thi Thu Nguyen was also responsible for the acquisition of Vietnamese native speakers for our study. She administered the collection of the LEAP-Q and contacted four native speakers in Vietnam to participate in the stimuli rating study.

Katharina Spalek is the principal investigator of the project "Focus alternatives in the human mind: Retrieval, Representation and Recall" (FAHMRRR) and initiated this study. In addition, she supervised the project and worked closely together with Huong Thi Thu Nguyen and Annika Tjuka while developing the study. Katharina Spalek also performed the final analysis of the data presented in this paper.

All authors worked closely together while writing the present article.

References

Anderson, V., Hiramoto, M., & **Wong, A.** (2007). Prosodic analysis of the interactional particle *ne* in women's and men's Japanese. *Japanese/Korean Linguistics, 15*, 43–54.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & **Plumb, I.** (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241–251. DOI: https://doi.org/10.1111/1469-7610.00715

Bates, D., Maechler, M., Bolker, B., & **Walker, S.** (2014). lme4: Linear mixed-effects models using eigen and s4. *R package version, 1*(7), 1–23. DOI: https://doi.org/10.18637/jss.v067.i01

Beaver, D. I., & **Clark, B. Z.** (2009). *Sense and sensitivity: How focus determines meaning* (Vol. 12). United Kingdom: John Wiley & Sons.

Bishop, J. (2016). Individual differences in top-down and bottom-up prominence perception. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel & N. Veilleux (Eds.), *Proceedings of Speech Prosody 2016* (pp. 668–672). Boston, USA. DOI: https://doi.org/10.21437/SpeechProsody.2016-137

Boersma, P., & **Weenink, D.** (2009). *Praat: Doing phonetics by computer (version 5.1.13).* Retrieved from http://www.praat.org

Braun, B., & **Tagliapietra, L.** (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes, 25*(7-9), 1024–1043. DOI: https://doi.org/10.1080/01690960903036836

Brunelle, M. (2009). Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics, 37*(1), 79–96. DOI: https://doi.org/10.1016/j.wocn.2008.09.003

Brunelle, M. (2017). Stress and phrasal prominence in tone languages: The case of Southern Vietnamese. *Journal of the International Phonetic Association, 47*(3), 283–320. DOI: https://doi.org/10.1017/S0025100316000402

Brunelle, M., Hạ, K. P., & **Grice, M.** (2012). Intonation in Northern Vietnamese. *The Linguistic Review, 29*, 3–36. DOI: https://doi.org/10.1515/tlr-2012-0002

Brunelle, M., Nguyễn, D. D., & **Nguyễn, K. H.** (2010). A laryngographic and laryngoscopic study of Northern Vietnamese tones. *Phonetica, 67*(3), 147–169. DOI: https://doi.org/10.1159/000321053

Campbell, G. L. (2003). *Concise compendium of the worlds languages.* Routledge. DOI: https://doi.org/10.4324/9780203018057

Đỗ, T. T., Trần, T. H., & **Boulakia, G.** (1998). Intonation in Vietnamese. In E. Garding (Ed.), *Intonation systems: A survey of twenty languages* (pp. 398–420). Cambridge, United Kingdom: Cambridge University Press.

Frank, C. K., Baron-Cohen, S., & **Ganzel, B. L.** (2015). Sex differences in the neural basis of false-belief and pragmatic language comprehension. *NeuroImage, 105*, 300–311. DOI: https://doi.org/10.1016/j.neuroimage.2014.09.041

Fraundorf, S. H., Watson, D. G., & **Benjamin, A. S.** (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language, 63*(3), 367–386. DOI: https://doi.org/10.1016/j.jml.2010.06.004

Freitag, C. M., Retz-Junginger, P., Retz, W., Seitz, C., Palmason, H., Meyer, J., … von Gontard, A. (2007). Evaluation der deutschen Version des Autismus-Spektrum-Quotienten (AQ) – die Kurzversion AQ-k. *Zeitschrift für Klinische Psychologie und Psychotherapie, 36*(4), 280–289. DOI: https://doi.org/10.1026/1616-3443.36.4.280

Fujisaki, H. (1988). A note on the physiological and physical basis for the phrase and accent components in the voice fundamental frequency contour. In O. Fujimura (Ed.), *Vocal physiology: Voice production, mechanisms and functions* (pp. 347–355). Raven Press.

Goldsmith, J. (1976). *Autosegmental phonology* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Gotzner, N., Wartenburger, I., & **Spalek, K.** (2016). The impact of focus particles on the recognition and rejection of contrastive alternatives. *Language and Cognition, 8*(1), 59–95. DOI: https://doi.org/10.1017/langcog.2015.25

Gussenhoven, C. (2004). *The phonology of tone and intonation.* Cambridge, United Kingdom: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511616983

Hạ, K. P. (2012). *Prosody in Vietnamese – Intonational form and function of short utterances in conversation*. Asia-Pacific Linguistics (SEAMLESS).

Herlitz, A., & **Rehnman, J.** (2008). Sex differences in episodic memory. *Current Directions in Psychological Science, 17*(1), 52–56. DOI: https://doi.org/10.1111/j.1467-8721.2008.00547.x

Hoàng, T. C. (1989). *Tiếng Việt trên các miền đất nước [Vietnamese in all regions of the country]*. Vietnam, Hanoi: Hanoi National University Publishing House.

Hole, D. (2008). EVEN, ALSO and ONLY in Vietnamese. In S. Ishihara, S. Petrova & A. Schwarz (Eds.), *Interdisciplinary Studies in Information Structure, 11*, 1–54. Potsdam, Germany: Universitaetsverlag Potsdam.

Hung, A.-Y., & **Cheng, Y.** (2014). Sex differences in preattentive perception of emotional voices and acoustic attributes. *Neuroreport, 25*(7), 464–469. DOI: https://doi.org/10.1097/WNR.0000000000000115

Husband, E. M., & **Ferreira, F.** (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition and Neuroscience, 31*(2), 217–235. DOI: https://doi.org/10.1080/23273798.2015.1083113

Ip, M. H. K., & **Cutler, A.** (2017). Intonation facilitates prediction of focus even in the presence of lexical tones. In *Proceedings of Interspeech 2017* (pp. 1218–1222). Stockholm, Sweden: International Speech Communication Association (ISCA). DOI: https://doi.org/10.21437/Interspeech.2017-264

Jannedy, S. (2007). Prosodic focus in Vietnamese. *Interdisciplinary Studies on Information Structure, 8*, 209–230.

Kaiser, A., Kuenzli, E., Zappatore, D., & **Nitsch, C.** (2007). On females' lateral and males' bilateral activation during language production: A fMRI study. *International Journal of Psychophysiology, 63*(2), 192–198. DOI: https://doi.org/10.1016/j.ijpsycho.2006.03.008

Kansaku, K., Yamaura, A., & **Kitazawa, S.** (2000). Sex differences in lateralization revealed in the posterior language areas. *Cerebral Cortex, 10*(9), 866–872. DOI: https://doi.org/10.1093/cercor/10.9.866

Koch, X., & **Spalek, K.** (in progress). Contrastive intonation effects on word recall for information-structural alternatives.

Krifka, M. (2006). Basic Notions of Information Structure. In C. Féry, G. Fanselow & M. Krifka (Eds.), *The Notions of Information Structure. Interdisciplinary studies on information structure* (pp. 13–56). Potsdam, Germany: Universitaetsverlag Potsdam.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge, United Kingdom: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511808814

Lee-Wong, S. M. (1994). Imperatives in requests: Direct or impolite – Observations from Chinese. *Pragmatics, 4*(4), 491–515. DOI: https://doi.org/10.1075/prag.4.4.01lee

Marian, V., Blumenfeld, H. K., & **Kaushanskaya, M.** (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967. (Vietnamese translation by Hien Phạm, University of Alberta, Canada and Quyen Nguyễn, Vietnam National University, Hanoi). DOI: https://doi.org/10.1044/1092-4388(2007/067)

Michaud, A. (2004). Final consonants and glottalization: New perspectives from Hanoi Vietnamese. *Phonetica, 61*(2–3), 119–146. DOI: https://doi.org/10.1159/000082560

Michaud, A., & **Brunelle, M.** (2016). *Information structure in Asia: Yongning Na (Sino-Tibetan) and Vietnamese (Austroasiatic)*. Oxford, United Kingdom: Oxford University Press.

Michaud, A., & **Vũ, T. N.** (2004). Glottalized and Nonglottalized Tones under Emphasis: Open Quotient Curves remain stable, F0 Curve is modified. In B. Bel & I. Marlien (Eds.), *Proceedings of Speech Prosody 2004* (pp. 745–748). Nara, Japan: s.n.

Miller, T., Athanasopoulou, A., Pincus, N., & **Vogel, I.** (2015). The Effect of Focus on Phonation in Northern Vietnamese Tones. *Poster presented at Linguistic Society of America.*

Nieuwland, M. S., Ditman, T., & **Kuperberg, G. R.** (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language, 63*(3), 324–346. DOI: https://doi.org/10.1016/j.jml.2010.06.005

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1*(1), 75–116. DOI: https://doi.org/10.1007/BF02342617

Schirmer, A., Kotz, S. A., & **Friederici, A. D.** (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14*(2), 228–233. DOI: https://doi.org/10.1016/S0926-6410(02)00108-8

Schirmer, A., Striano, T., & **Friederici, A. D.** (2005). Sex differences in the preattentive processing of vocal emotional expressions. *NeuroReport, 16*(6), 635–639. DOI: https://doi.org/10.1097/00001756-200504250-00024

Schirmer, A., Zysset, S., Kotz, S. A., & **von Cramon, D. Y.** (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *NeuroImage, 21*(3), 1114–1123. DOI: https://doi.org/10.1016/j.neuroimage.2003.10.048

Spalek, K., Gotzner, N., & **Wartenburger, I.** (2014). Not only the apples: Focus sensitive particles improve memory for information-structural alternatives. *Journal of Memory and Language, 70,* 68–84. DOI: https://doi.org/10.1016/j.jml.2013.09.001

Thompson, L. C. (1965). *A Vietnamese Grammar.* ERIC.

Vũ, T. P. (1982). Phonetic Properties of Vietnamese Tones across Dialects. *Papers in Southeast Asian Linguistics,* 55–75.

Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., & **Grodd, W.** (2002). Dynamic brain activation during processing of emotional intonation: Influence of acoustic parameters, emotional valence, and sex. *NeuroImage, 15*(4), 856–869. DOI: https://doi.org/10.1006/nimg.2001.0998

Xiang, M., Grove, J., & **Giannakidou, A.** (2013). Dependency-dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology, 4,* 1–19. DOI: https://doi.org/10.3389/fpsyg.2013.00708

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics, 27*(1), 55–105. DOI: https://doi.org/10.1006/jpho.1999.0086

Yan, M., & **Calhoun, S.** (2019). Priming Effects of Focus in Mandarin Chinese. *Frontiers in Psychology, 10,* 1–16. DOI: https://doi.org/10.3389/fpsyg.2019.01985