Intonation guides sentence processing in the left inferior frontal gyrus

Constantijn L van der Burght[a†], Tomás Goucha[a†], Angela D Friederici[a], Jens Kreitewolf[a,b]*, Gesa Hartwigsen[a]*

[a] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

[b] Department of Psychology, University of Lübeck, 23562 Lübeck, Germany

[†,]* Authors contributed equally

Declarations of interest:

Corresponding author:
Constantijn van der Burght
Department of Neuropsychology
Max Planck Institute for Human Cognitive and Brain Sciences
Stephanstr. 1a
04103 Leipzig, Germany
phone: +49 (0)341 9940 2238
email: vanderburght@cbs.mpg.de

Abstract

Speech prosody, the variation in sentence melody and rhythm, plays a crucial role in sentence comprehension. Specifically, changes in intonational pitch along a sentence can affect our understanding of who did what to whom. To date, it remains unclear how the brain processes this particular use of intonation and which brain regions are involved. In particular, one central matter of debate concerns the lateralisation of intonation processing. To study the role of intonation in sentence comprehension, we designed a functional MRI experiment in which participants listened to spoken sentences. Critically, the interpretation of these sentences depended on either intonational or grammatical cues. Our results showed stronger functional activity in the left inferior frontal gyrus (IFG) when the intonational cue was crucial for sentence comprehension compared to when it was not. When instead a grammatical cue was crucial for sentence comprehension, we found involvement of an overlapping region in the left IFG, as well as in a posterior temporal region. A further analysis revealed that the lateralisation of intonation processing depends on its role in syntactic processing: activity in the IFG was lateralised to the left hemisphere when intonation was the only source of information to comprehend the sentence. In contrast, activity in the IFG was right-lateralised when intonation did not contribute to sentence comprehension. Together, these results emphasise the key role of the left IFG in sentence comprehension, showing the importance of this region when intonation establishes sentence structure. Furthermore, our results provide evidence for the theory that the lateralisation of prosodic processing is modulated by its linguistic role.

1. Introduction

In everyday conversation, different components of speech need to be integrated to understand the meaning of what someone is saying. Besides the meaning of each individual word in the sentence, additional information is crucial to understand *who did what to whom* (Everaert, Huybregts, Chomsky, Berwick, & Bolhuis, 2015; Sportiche, Koopman, & Stabler, 2013). Besides syntax, one way in which this information can be conveyed is through prosody: the changes in rhythm and melody of speech (Cutler, Dahan, & van Donselaar, 1997). An important feature of prosody is intonation, marked by the changes in pitch along a sentence. Intonation can dramatically change the interpretation of a sentence. For instance, the sentence "the teacher said the student is mistaken" has two possible interpretations depending on the particular intonation. One can either say "the teacher said: the student is mistaken" or "the teacher, said the student, is mistaken". Here, the particular use of intonation determines whether the teacher or the student is alleged to be wrong, by creating boundaries between different parts of the sentence. The prosodic features marking these boundaries are a pause between the two sentence parts, preceded by a rise in pitch and a lengthening of the syllable before the pause. Together, these features acoustically separate two parts of a sentence, and constitute a so-called *intonational phrase boundary* (IPB) (Selkirk, 1984).

Despite many years of neurocognitive research on prosody, it remains largely unknown how exactly intonation contributes to sentence comprehension, and what the brain implementation of this process is. Although pitch was shown to be preferably processed in the right hemisphere (Zatorre, 2001), early neurocognitive models on prosody postulated that the stronger the linguistic function of prosody the larger the leftward lateralisation (Friederici & Alter, 2004; Van Lancker, 1980). A recent study using intracranial cortical recordings showed that intonation is processed in specific neural populations in the temporal lobe that are not involved in processing other speech components such as the sounds of words (Tang, Hamilton, & Chang, 2017). Yet, intonation is rapidly integrated with other phonetic components (e.g. consonants and vowels) to interpret a sentence, as has been shown in early behavioural studies (Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992). Furthermore, electrophysiological studies have demonstrated that prosodic information and information about sentence structure are integrated online during sentence processing (Friederici, Cramon, & Kotz, 2007; Männel & Friederici, 2009; Sammler, Kotz,

Eckstein, Ott, & Friederici, 2010; Steinhauer, Alter, & Friederici, 1999, see Bögels et al. (2011) for a review). To date, neuroimaging studies have been inconclusive with regard to the brain regions involved in this use of intonation for sentence comprehension. Functional neuroimaging research has focused on particular parts of prosody rather than its actual contribution to sentence comprehension. For instance, a network of superior temporal and fronto-opercular regions in the right hemisphere has been found to be involved in speech processing depending on the presence of pitch information (Meyer, Alter, Friederici, Lohmann, & Cramon, 2002; Meyer, Steinhauer, Alter, Friederici, & Cramon, 2004; Plante, Creusere, & Sabin, 2002). A similar fronto-temporal network has been found in the perception of natural compared to hummed speech (Ischebeck, Friederici, & Alter, 2008). However, since these previous studies compared various types of filtered speech, they focused on the acoustic processing of intonation rather than its use for sentence comprehension. Only few neuroimaging studies have investigated which brain regions are involved when prosodic information guides sentence comprehension. These studies either involved a rather quantitative analysis of the intonational cue (e.g., the presence of two intonational phrase boundaries versus a single one (Ischebeck et al., 2008)) or compared conditions in which the stimuli were not lexically matched (Strelnikov, Vorobyev, Chernigovskaya, & Medvedev, 2006). Consequently, several brain regions have been found to support the processing of intonational contours in speech, but it is unknown whether these regions also play a role in guiding sentence comprehension.

Aside from intonation, grammatical cues can guide sentence comprehension, by means of a particular word form (morphosyntax). For example, in the sentence "The teachers said the student is mistaken" the word form of *teachers* and *said* (both signal plural) and *student* and *is* (both singular) establishes that the teachers describe the student's behaviour, not the other way around. This example shows how the sentence structure can be established by grammatical cues. Previous work has shown that these grammatical cues are processed in a left-hemisphere network of frontal and temporal regions (see Friederici (2011) and Hagoort (2014) for reviews). Specifically, when grammatical cues are the only informative elements available to interpret a sentence structure, the posterior part of the left inferior frontal gyrus (IFG) has been shown to be engaged (Goucha & Friederici, 2015). Functional imaging and lesion studies have further shown that successful processing of grammatical cues relies on an intact left superior temporal cortex (Bornkessel, Zysset,

Friederici, Cramon, & Schlesewsky, 2005; Regel, Kotz, Henseler, & Friederici, 2017; Rolheiser, Stamatakis, & Tyler, 2011).

Taken together, grammatical as well as intonational cues can guide sentence comprehension by resolving ambiguities and establishing the structure of a sentence. Whereas the cortical network supporting processing of grammatical cues has been extensively studied, it remains poorly understood how exactly intonation contributes to sentence comprehension, and what the neural correlate of this contribution is. We aimed to fill this gap by studying processing of spoken sentences in which either intonational cues or grammatical cues are fundamental to understand what is being said. We designed a functional magnetic resonance imaging (fMRI) paradigm to achieve two goals. The first goal was to investigate the contribution of intonational and grammatical cues for sentence comprehension. The second goal was to study the hemispheric lateralisation of intonation processing. To achieve this, participants had to comprehend specific sentence types (see Figure 1). Across conditions, the sentence structure was established by different language cues: sentences could be interpreted in two possible ways until a point at which the cue ensured only one possible interpretation (Marslen-Wilson et al., 1992). This cue was either intonational (an IPB) or grammatical. The grammatical cue was established by morphological case marking of a personal pronoun, such that it matched only one of the two verbs in the sentence. The paradigm centred around the following sentence, which is open to two interpretations:

Peter verspricht Nick dafür zu bezahlen

(i) *Peter **promises Nick** to pay for it*

(ii) *Peter promises **to pay Nick** for it*

In these sentences, intonational and/or grammatical cues are required to convey who did what to whom. Without them, listeners cannot identify whether *Nick* was promised something or paid for something instead. In our key conditions, the position of an intonational phrase boundary (marked with #) helped the listener to identify one of the two possible interpretations:

(A) Peter verspricht Nick # dafür zu bezahlen

*Peter **promises Nick** to pay for it*

(B) Peter verspricht # Nick dafür zu bezahlen

*Peter promises **to pay Nick** for it*

5

Alternatively, the presence of a grammatical cue established a clear interpretation. Here, we made use of German case marking, which distinguishes between the different roles a word can have in a sentence. Depending on the word form and its corresponding case (dative or accusative), the personal pronoun can only be the object of the verb *to promise* or *to pay*.

> (C) Peter verspricht sie dafür zu bezahlen
>
> *Peter promises to pay^{+ACC} her^{ACC} for it*

As a second goal of our study, we investigated the hemispheric lateralisation of intonation processing. Although a general consensus exists that processing of linguistic components such as grammatical cues is predominantly left-lateralised (Friederici, 2011; Vigneau et al., 2006), it has long been debated whether a similar lateralisation exists for the processing of intonation (Luks, Nusbaum, & Levy, 1998; van Lancker, 1980; Wildgruber et al., 2004). In general, the right hemisphere is seen as dominant in the processing of pitch information, including intonation (Poeppel, 2003). However, dichotic listening (Luks et al., 1998) and lesion studies (reviewed in Witteman et al. (2011)) have shown that lateralisation depends on the linguistic function of intonation. Moreover, results from functional neuroimaging studies suggest that intonation processing in the IFG and temporal cortex is lateralised, but the contribution of either hemisphere depends on the specific control task used (Kreitewolf, Friederici, & Kriegstein, 2014). It is likely that processing of linguistic prosody relies on fronto-temporal networks in both hemispheres (Belyk & Brown, 2014; Witteman et al., 2011), with a dominance of the left hemisphere when pitch information is used to signal linguistic aspects (Friederici & Alter, 2004; van Lancker, 1980). However, it remains elusive whether intonation processing is lateralised when it contributes to sentence comprehension, and to which hemisphere. We addressed this question by investigating the neural processing of intonation, focusing specifically on its use for sentence comprehension. Our paradigm allowed for an investigation of the linguistic importance of prosody, since we varied the linguistic role of the intonational cue across conditions while keeping acoustical information identical. To this end, we included a condition in which the IPB was present but not essential to establish the sentence structure (see Figure 1).

In summary, our study was designed to answer two separate questions. First, we investigated how the presence of intonational and grammatical cues influences sentence processing. We hypothesised that processing depends on the availability of the specific cue type in the sentence, and whether this cue appeared in isolation or in combination with a

6

second cue. Based on previous studies (Goucha & Friederici, 2015; Kreitewolf et al., 2014), we expected increased activity in the left IFG (and possibly additional bilateral superior temporal regions) in conditions in which intonation was the only cue establishing the structure. When sentence comprehension relied on a grammatical cue only, we also expected increased activity in the left IFG, possibly with additional recruitment of the left posterior temporal cortex (Bornkessel et al., 2005; Regel et al., 2017). Our second research question concerned the lateralisation of intonation processing. Specifically, we investigated whether the role of the intonational cue in establishing the sentence structure determined the lateralisation of brain areas involved in intonation processing. Based on previous work (e.g Kreitewolf et al., 2014), we hypothesised that lateralisation depends on the linguistic function of the intonational cue. For processing of intonational cues that establish sentence structure, we expected left-lateralised activity of core language regions (IFG and posterior superior temporal gyrus (pSTG)). In contrast, we expected a shift towards the right IFG when prosodic content was present, but not used to establish the structure of the sentence.

2. Methods

2.1. *Participants*

Twenty-six native German speakers (15 female; mean age: 26.3 years; age range: 20-33 years) were included in the final analyses. All participants were right-handed (Oldfield, 1971) and had normal or corrected-to-normal vision. All reported normal hearing and none were professional musicians. None had a history of neurological or psychiatric illness, drug or alcohol abuse, chronic medical disease, or any other contraindication against participation in an MRI experiment. Twelve additional participants had to be excluded because they did not complete the experiment (n = 2) or because they performed below chance level in at least one of the six stimulus conditions (n = 10). The sample size was determined based on previous fMRI studies on sentence processing (e.g. Goucha & Friederici, 2015; Kristensen, Wang, Petersson, & Hagoort, 2013; Perrone-Bertolotti, Dohen, Lœvenbruck, Sato, Pichat & Baciu 2013). The exclusion criteria were established prior to data analysis. All participants gave written consent prior to participating in the experiment, which was approved by the ethics committee of the University of Leipzig.

2.2. *Experimental design*

To investigate the effect of intonational and grammatical cues on sentence comprehension, we used an event-related fMRI design that employed six different sentence types with varying amount of intonational and grammatical information (see Figure 1A). The stimulus set was built around the following sentence:

A verspricht B dafür zu bezahlen (literally: *A promises B for it to pay*)

which has two possible interpretations.

    i.   [A verspricht B [dafür zu bezahlen]]

        [A promises B [to pay for it]]

    ii.   [A verspricht [B dafür zu bezahlen]]

        [A promises [to pay B for it]

In German, the two structures (i) and (ii) can be realised by an identical string of words. This requires specific language cues to distinguish the two possible meanings and to clarify whether *B* is the object of the verb *to promise* or *to pay*. One such cue is the position of an IPB (indicated with "#"), which can create the distinction as follows:

Prosody Only 1 (1)   [Peter verspricht Nick # [dafür zu bezahlen]]

(*ProsOnly1*)   *[Peter promises Nick # [to pay for it]]*

Prosody Only 2 (2)   [Peter verspricht # [Nick dafür zu bezahlen]]

(*ProsOnly2*)   *[Peter promises # [to pay Nick for it]]*

The IPB acoustically divides the sentence and groups the proper noun *Nick* to either of the two verbs. Without the IPB, *ProsOnly1* and *ProsOnly2* are ambiguous. The IPB is defined by a pitch rise and syllable lengthening, followed by a pause (Selkirk, 1984) (see Figure 1B).

An additional cue can resolve the ambiguity, for example, when a personal pronoun is used (such as *she*) instead of a proper noun (*Nick*). In German, personal pronouns are inflected, meaning that their morphosyntactic form defines their role in the sentence (i.e., by case marking). Similarly, verbs require objects in a specific case. For example, "versprechen" *(to promise)* requires objects in the dative case, whereas "bezahlen" *(to pay)* requires an accusative. Making use of the German case marking system, we constructed sentences in which the structure is built by a grammatical cue only:

Grammatical Only (3)   [Peter verspricht [sie dafür zu bezahlen]]

(*GramOnly*)   *[Peter promises^{+DAT} [to pay^{+ACC} her^{ACC} for it]]*

Although the position of the word "sie" (*her*) does not yet clarify to what verb the word belongs, the case marking of the word ensures that "sie" is necessarily an object of *to pay* and cannot belong to the verb *to promise.* The sentence can only be interpreted in one way because of the morphosyntactic form of *her^{ACC}* ("sie").

To investigate sentence processing guided by these cues, we designed control conditions with additional cues, for example:

Baseline Prosody Only (4)   [Peter verspricht Nick # [sie zu bezahlen]]

(*BL ProsOnly*)   *[Peter promises^{+DAT} Nick # [to pay^{+ACC} her^{ACC} for it]]*

In this sentence, identification of the syntactic structure is facilitated by the additional grammatical cue "sie", as compared to (*ProsOnly1*). Thus, it is not necessary to disambiguate the verb-argument structure because two objects (*Nick* and *her*) are present in this sentence.

Similarly, as a control condition for the experimental condition *GramOnly*, we created sentences that contained an intonational cue in addition to the grammatical cue, for example:

Baseline Grammatical Only (5)   [Peter verspricht # [sie dafür zu bezahlen]]

(*BL GramOnly*)   *[Peter promises^{+DAT} # [to pay^{+ACC} her^{ACC} for it]]*

A final baseline condition was created, in which intonation was not required to understand who did what to whom the in the sentence. This sentence type had an IPB, as in *ProsOnly1* and *ProsOnly2,* but the verb-argument structure did not have to be resolved.

Baseline Prosody No Choice (6)   [Peter verspricht # [heute dafür zu bezahlen]]
(*BL ProsNoChoice*)   *[Peter promises^{+DAT} # [to pay^{+ACC} for it today]]*

Each condition consisted of 50 unique verb combinations matched to a variety of German first names (yielding a total of 300 sentences). In each sentence, the verb in the main clause required an object in the dative form and the verb in the subordinate clause

required an object in the accusative form, or vice versa. This ensured that the grammatical cue (the personal pronoun), in either the dative or accusative form, could be unambiguously assigned to either verb. The matching of dative and accusative verbs to main clause or subordinate clause was pseudo-randomised across trials.

To confirm that participants were equally likely to attribute the object in the sentence to the first or the second verb, we calculated if participants had a response bias for either condition *ProsOnly1* or *ProsOnly2*. Using methods of signal detection theory (Macmillan & Creelman, 1991) this response bias turned out not to be significant from 0, suggesting that participants had no intrinsic bias for either syntactic structure (see Meyer et al., 2016, for a similar approach).
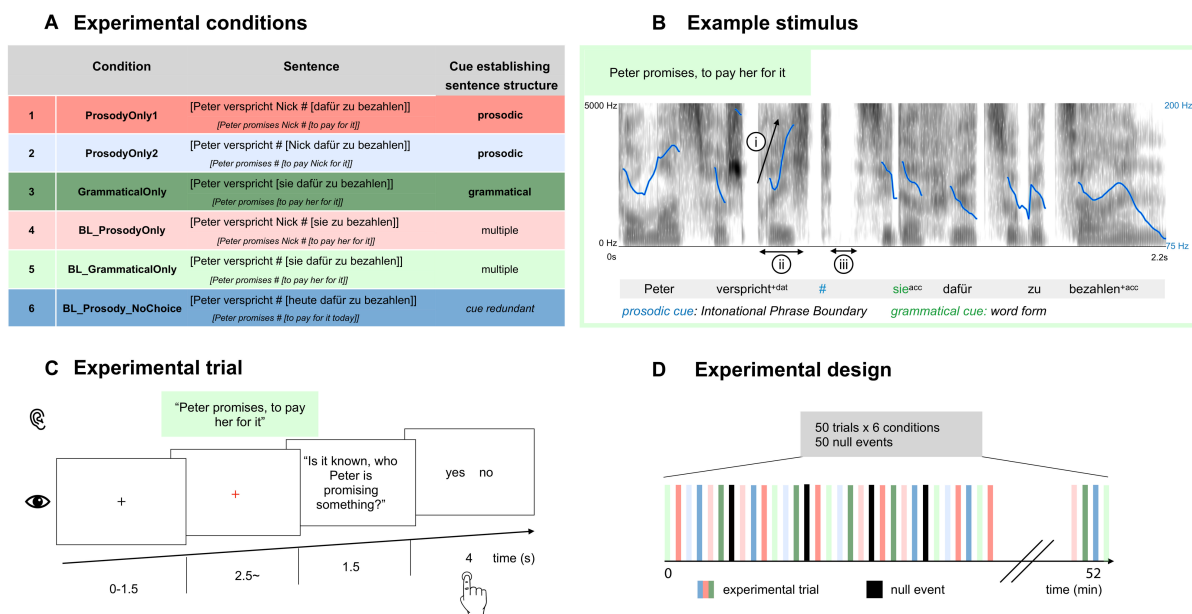
Figure 1. Experimental design. (A) Experimental conditions. Across conditions, a different combination of language cues established the verb-argument structure; that is, whether the object in the sentence ("Nick" or "she") belongs to the first or second verb. Prosodic cue (indicated by "#"): an acoustic break marking a transition in a sentence (i.e., intonational phrase boundary; IPB). Grammatical cue (in bold typeface): morphosyntactic case marking of the personal pronoun "sie" (*she*), matched to either the verb in the main clause ("verspricht" (*promises*)) or sub clause ("zu bezahlen" (*to pay*)). 'Prosody': presence of a prosodic cue. 'Grammatical': presence of a grammatical cue. 'Only' indicates that that cue was the only cue present in that sentence. 'BL': baseline condition matched to the specific cue. 'NoChoice': the prosodic cue did not influence the response choice in the task. **(B)** Spectrogram of an example stimulus with intonational and grammatical cues. Here, the syntactic structure is established by both the grammatical cue "sie" (*she*) and an intonational cue in shape of an IPB. The IPB is composed of three acoustic events, indicated with arrows: (i) a pitch rise and (ii) lengthening of the syllable, followed by (iii) a pause. **(C)** Overview of an experimental trial. **(D)** Overview of the fMRI session.

2.3. Experimental procedure

Each experimental trial (see Figure 1C) started with a white fixation cross, which turned red 200 ms prior to auditory stimulation to alert the participant. Subsequently, the spoken sentence was presented. After each sentence, a visual comprehension question was shown for 1500 ms, to which participants were asked to respond with a right-handed button press. Participants had to respond within 4 s. The subsequent trial started after the response phase and an additional delay of 0, 750 or 1500 ms (uniformly jittered).

To ensure active listening, the comprehension question was visually presented after the sentence. The question probed sentence comprehension by asking whether an object was present in either the main or subordinate clause: participants were asked *Do you know who Peter promises something?* ("Weiß man, wem Peter etwas verspricht?") or *Do you know, who is paid?* ("Weiß man, wer bezahlt wird?"). The question types were presented pseudo-randomly, with each of the two question types occurring evenly across sentence conditions.

During the experimental session (Figure 1D), the six experimental conditions were presented in a pseudo-random order: two stimuli of the same condition were always separated by at least two trials so that the conditions were distributed evenly throughout the experiment. 50 null trials with an average-trial duration of 8.5 s were pseudo-randomly interspersed with the other conditions throughout the experiment. The experiment was performed within one session with a total duration of 52 minutes. Each participant performed a short practice session immediately before the fMRI experiment, which mirrored the main experiment but consisted of different stimuli.

Auditory stimuli were presented through MR-compatible headphones (MR confon OPTIME 1; MR confon GmbH, Magdeburg, Germany). Participants additionally wore earplugs to attenuate scanner noise. Stimulus presentation and response collection were controlled via Presentation (Neurobehavioural Systems, Inc., Albany, CA, USA), with visual stimuli presented on an LCD projector (PLC-XP50L, SANYO, Tokyo, Japan). Participants could see the projection via a mirror that was attached to the head coil.

2.4. Stimulus properties

Sentences were spoken by a male, professional native German speaker and recorded in a sound-attenuating chamber (IAC – I200 series, Winchester, United Kingdom). The

11

digitised speech signals (sampling rate 44.1kHz; resolution 16 bits) were adjusted to the same root mean square value using MATLAB (Mathworks, Inc., Sherborn, MA, USA). To ensure consistent comparisons between stimuli with matching main or subordinate clauses, all stimuli were cross-spliced: the clauses before and after the pause were cut in Adobe® Audition™ CS5.5 and concatenated to form the stimulus sentences. This procedure ensured that identical sentence parts across conditions originated from the same recording.

Importantly, in this way we guaranteed that in contrasts between two sentences with IPBs, those IPBs were acoustically identical. That is, for *ProsOnly1* & *BL_ProsOnly*, and for *ProsOnly2*, *BL_GramOnly* & *BL_ProsNoChoice*, the onsets of the stimuli up to and including the IPB originated from the same recordings.

To further improve acoustic consistency across the stimulus set, in those sentences that contained an IPB (all except *GramOnly*) we introduced a pause of constant duration (100 ms): all first parts of the stimuli were cut *until* the pause, to which a pause of constant duration was added, followed finally by the second part of the stimuli (which had been cut *after* the pause). We chose 100 ms based on pilot study results, showing that such a pause could be clearly perceived and sounded natural.

The *GramOnly* condition, containing no IPB, was also constructed by cross-splicing two elements. To prevent the realisation of an IPB in the first element, we had the speaker produce a sentence without a syntactic boundary after the verb (where an IPB would be illegal). Subsequently, we spliced the recording after the verb and concatenated it to the same sentence ending as in *BL_GramOnly*. This yielded a sentence with natural prosody but without any of the three acoustic cues characterising the IPB (pitch rise, syllable lengthening and a pause). Furthermore, this ensured that the sentence endings of *GramOnly* and its baseline equivalent were matched. Spectrograms of all 6 sentence conditions are provided in Supplementary Figure 1.

Participant debriefings and a pilot study on a separate sample of participants (n = 18) confirmed that all stimuli were perceived as natural, grammatical, and non-ambiguous.

2.5. fMRI acquisition

Functional imaging was performed on a 3 Tesla Siemens Skyra scanner (Siemens Healthcare, Erlangen, Germany) using a 20-channel head coil. A gradient-echo echo-planar-imaging (EPI) sequence was run (acquisition time [TA] = 2s; continuous scanning; echo delay time [TE] =

30ms; flip angle 78°; matrix size 64 × 64; field of view [FOV] = 192 × 192mm; 30 slices of 3mm thickness; in-plane resolution = 3mm x 3mm; gap = 1mm). For anatomical registration, T1-weighted images were either acquired during the scanning session or obtained from the in-house database when available.

2.6. Data analysis

2.6.1. Behavioural data

Response times and accuracy data were analysed using SPSS (IBM Corp., Armonk, NY, USA). Since behavioural measures were not normally distributed (Kolmogorov-Smirnov tests; all $p <$ 0.05), Friedman tests were used as a non-parametric alternative to repeated-measures analyses of variance. Follow-up Wilcoxon signed-rank tests were performed as post-hoc tests. Initial p-values lower than $\alpha = 0.05$ (two-tailed) were considered significant for all comparisons. To correct for multiple comparisons (a total number of 15), Bonferroni corrections were applied, yielding a corrected α-level of 0.0033 (0.05/15).

2.6.2. fMRI data

fMRI data were pre-processed and statistically analysed using SPM12 (www.fil.ion.ucl.ac.uk/spm, Wellcome Trust Centre for Neuroimaging). All functional images were realigned to the first image in the time series to correct head motion and unwarped to correct distortions caused by inhomogeneity in the magnetic field. After the T1-weighted image was co-registered to the mean EPI image, it was normalised to the Montreal Neurological Institute (MNI) template image. The deformation parameters resulting from this step were used to normalise all EPI images to MNI space. Finally, the data were smoothed using an isotropic Gaussian kernel of 8mm full-width at half-maximum.

Statistical analysis of the fMRI data was performed using a general linear model in SPM12. The onset and duration of each sentence were modelled per condition and convolved with the canonical hemodynamic response function. To account for domain-general effects of task performance on brain activation, we took into consideration between-condition differences in reaction times in our model. To this end, we built a regressor with response onsets and response times for each trial. This regressor was orthogonalized to the condition regressors and included in the general linear model (following Grinband, Wager, Lindquist, Ferrera, & Hirsch, 2008). Incorrect trials were

modelled as a separate noise condition. A high-pass filter with 128s cut-off was applied.

Contrast images of each condition of interest and participant were combined in a group

random effects analysis with a factorial design: the six experimental conditions entered as

levels of the factor CONDITION. Results were thresholded at an FWE-corrected cluster level

of $p < 0.05$, using an initial uncorrected voxel-wise threshold of $p < 0.001$ (Friston, Worsley,

Frackowiak, Mazziotta, & Evans, 1994). All activation peak coordinates are reported in MNI

space and the SPM anatomy toolbox (version 2.2c) (Eickhoff et al., 2005) was used for

anatomical localisation. Results were visualised using the BrainNet Viewer (Xia, Wang, & He,

2013).

Additionally, we performed a lateralisation analysis. This analysis was conducted by

normalizing the raw EPI images to a symmetrical MNI template. The first-level analysis was

run as described above, and the resulting contrast images were left-right flipped (Bozic,

Tyler, Ives, Randall, & Marslen-Wilson, 2010; Josse, Kherif, Flandin, Seghier, & Price, 2009;

Liégeois et al., 2002). On the second level, paired t-tests were run to compare the image of a

particular contrast of interest to its left-right flipped equivalent. We applied the same

statistical thresholds that were used in the activation analysis.

2.7. Availability of study materials

Data from this study have not been publicly archived since the conditions of our ethics

approval do not permit to do so. Analysis code and stimulus materials are available at

https://github.com/CLvanderBurght/prossyn/. No part of the study procedures or analyses

was pre-registered prior to the research being conducted.

3. Results

3.1. *Behaviour*

Both reaction times (RTs; Figure 2A) and accuracy rates (Figure 2B) differed across conditions, as shown by Friedman tests (RTs: $\chi^2(5) = 75.87$, $p < 0.001$; accuracy: $\chi^2(5) = 50.41$, $p < 0.001$). Pair-wise comparisons between conditions showed differences in the difficulty of sentence comprehension reflected in accuracy and reaction times. In general, participants showed lower accuracies and higher reaction times in conditions where only one cue was available compared to the matched baseline conditions. Accuracy decreased and reaction times increased in *ProsOnly1* (1) as compared to *BL_ProsOnly* (4), indicating that sentence comprehension was more difficult when only one language cue was present (RTs: $Z = -4.457$, $p < 0.001$; accuracy: $Z = -3.523$, $p < 0.001$). Similarly, sentence comprehension was more difficult in *GramOnly* (3), which contained only a grammatical cue, compared to *BL_GramOnly* (5), which contained both a grammatical and an intonational cue (RTs: $Z = -4.026$, $p < 0.001$; accuracy: $Z = -3.760$, $p < 0.001$). Other comparisons between matched sentences were significant (e.g. conditions 2 vs 6: RTs: $Z = -4.178$, $p < 0.001$; accuracy: $Z = -3.816$, $p < 0.001$; see supplementary tables S1 and S2 for complete pair-wise comparisons). Comparisons between matched conditions of interest (1 versus 2) or control conditions (4 versus 6) were not significant.
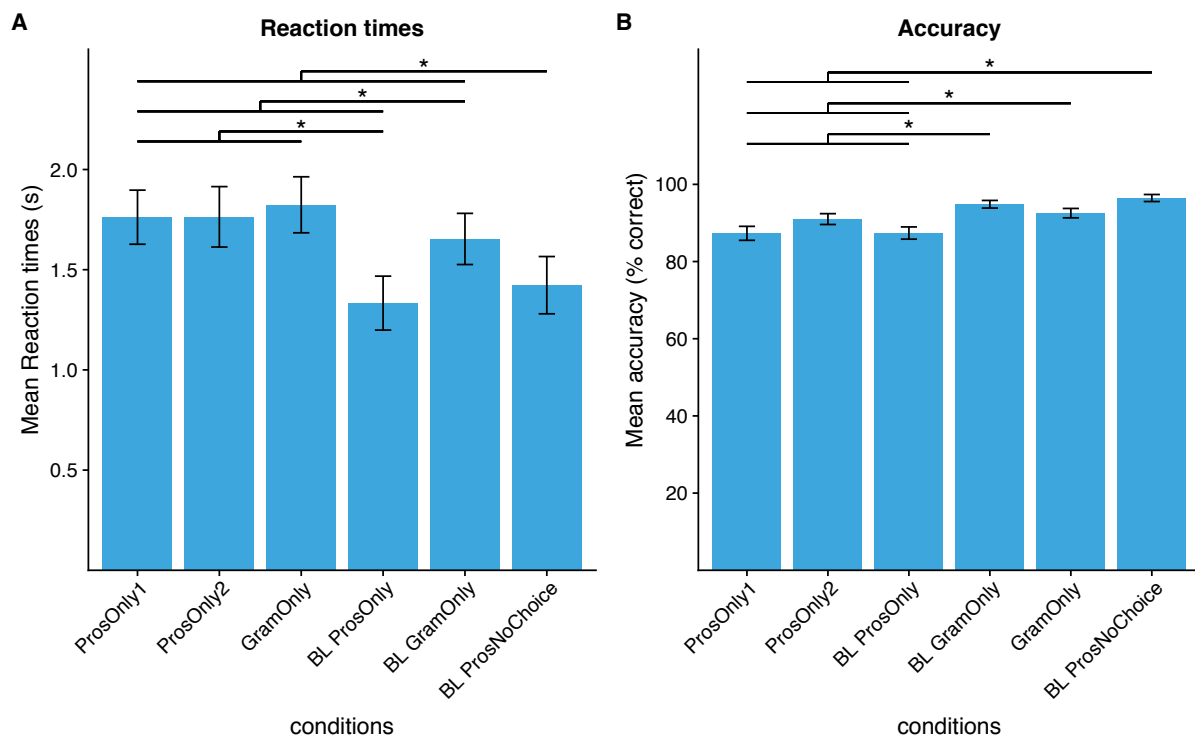
Figure 2: Behavioural results: reaction times (**A**) and accuracy rates (**B**) per condition. Error bars indicate ± 1 SEM. Lines between conditions and asterisks indicate pair-wise

comparisons with *p*-values smaller than 0.0033 (Bonferroni-corrected). BL stands for baseline.

3.2. fMRI – activity analysis

We investigated how the presence of specific language cues shapes sentence comprehension in the brain. To this end, we contrasted sentences in which only one specific language cue established the sentence structure to matched control conditions (i.e., 1 vs 4 and 3 vs 5, cf. Figure 1A). In the control sentences, processing of the sentence structure was facilitated in comparison to the conditions of interest, because multiple cues instead of a single cue marked the sentence structure. For an overview of significant activation clusters, see Table 1.

Syntactic structure established by a prosodic cue

To investigate processing of sentence structure established by prosodic information, the following experimental conditions were contrasted: *ProsOnly1* (1), in which a prosodic cue is the only factor establishing the sentence structure, versus *BL_ProsOnly* (4), in which additional cues determined the sentence structure (a grammatical cue). This contrast, reflecting sentence processing guided by a prosodic cue, yielded increased task-related activity in the left inferior frontal gyrus (IFG, peak activity at pars triangularis, x, y, z = -45, 29, -4; T = 4.93; Figure 3A and Table 1).
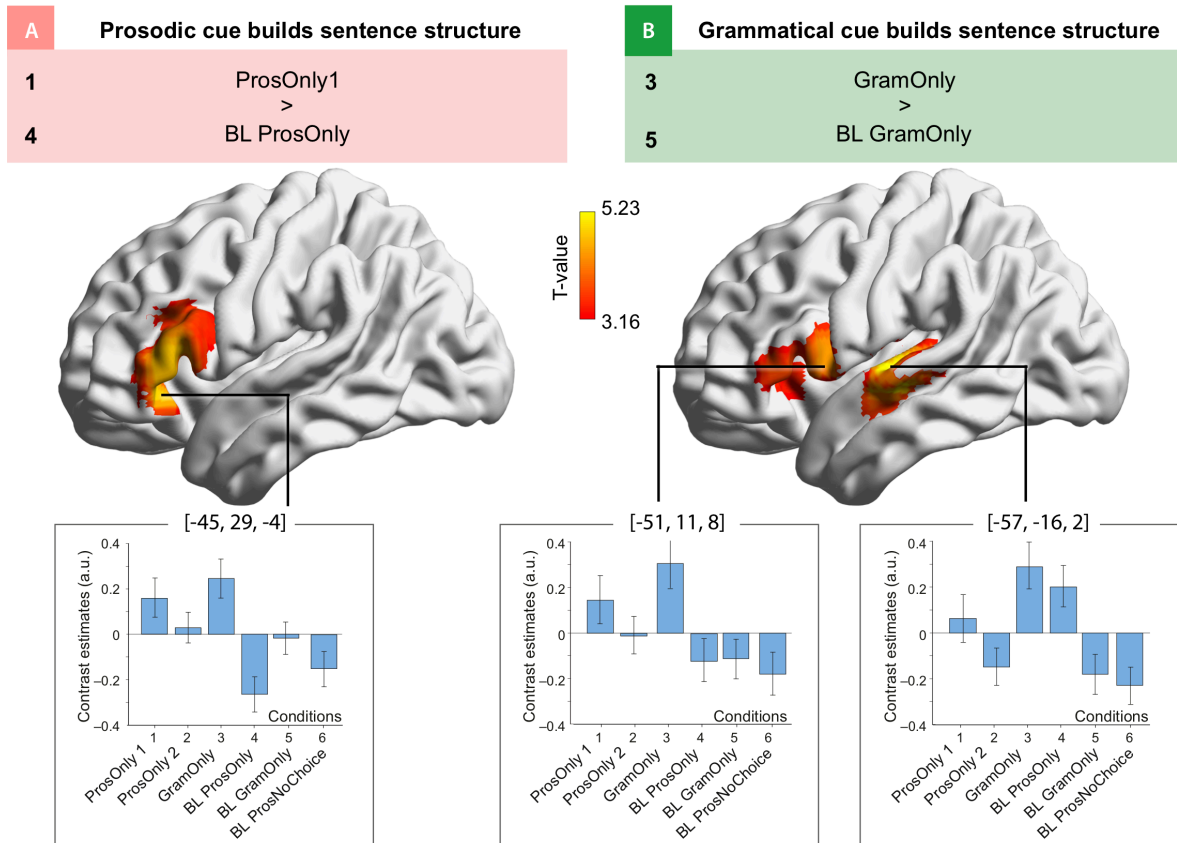
Syntactic structure established by a grammatical cue

To investigate processing of sentence structure marked by a grammatical cue, we contrasted *GramOnly* (3), in which only a grammatical cue marks the sentence structure, to a matched control condition in which an additional intonational cue establishes the sentence structure. Since in *BL_GramOnly* (5) the sentence structure was already established by the intonational cue, the grammatical cue was less important for resolving the structure. This contrast resulted in functional activation clusters in the pars opercularis of the left IFG (x, y, z = -51, 11, 8; T = 4.43) and the left superior temporal gyrus and sulcus (x, y, z = -57, -16, 2; T = 5.23; Figure 3B).

The reverse of the above described contrasts (3 > 1 and 4 > 5) did not yield significant activation clusters. Additionally, other contrasts between matched sentences (ProsOnly2 >

16

BL_GramOnly and ProsOnly2 > BL_ProsNoChoice) did not show significant results at *p* < 0.05, FWE-corrected (cluster-level).

Figure 3
Functional MRI results showing significant activation clusters for the different contrasts of interest. **(A)** Syntactic structure processing guided by a prosodic cue. **(B)** Syntactic structure processing guided by a grammatical cue. Bar plots show contrast estimates for each condition at activation peaks, indicated by the pointer, in arbitrary units (a.u.). All comparisons are thresholded on the cluster level at *p* < 0.05, FWE-corrected. Peak activity coordinates are in MNI space. BL stands for baseline.

3.3. fMRI – lateralisation analysis

A right-hemispheric dominance for intonation processing is often found in prosody research (Meyer et al., 2002; Sammler, Grosbras, Anwander, Bestelmeyer, & Belin, 2015). However, meta-analyses on prosody studies point towards a bilateral network for prosody processing (Belyk & Brown, 2014; Witteman et al., 2011). It has previously been suggested that intonation processing is left-lateralised specifically when pitch information is linguistically relevant (Friederici & Alter, 2004; Kreitewolf et al., 2014; van Lancker, 1980). However, this has not yet been shown with well-matched sentence stimuli. We therefore investigated

sentence conditions in which the intonation was matched acoustically but differed in terms of linguistic importance. Specifically, we compared conditions in which intonation guided sentence comprehension with matched conditions in which intonation was superfluous for sentence comprehension. The resulting contrast images were compared to their equivalent images in right-left flipped orientation. Results are summarised in Table 2.

First, we investigated lateralisation of intonation processing when the prosodic cue was crucial for sentence comprehension, assessed by the contrast *ProsOnly1* vs *BL_ProsOnly* (same contrast as in the activity analysis in Figure 3A). The results showed that the functional activation in the IFG was left-lateralised (x, y, z = -54, 29, 5, pars triangularis; T = 6.08; Figure 4A). Other areas that showed left-lateralised activity were the supplementary motor area (x, y, z = -6, 23, 50; T = 4.87) and the superior temporal gyrus (x, y, z = -51, -34, 2; T = 4.20). Additionally, functional activation was right-lateralised in the pre- and post-central gyrus (x, y, z = 30, -19, 56; T = 5.53) and in the superior temporal gyrus (x, y, z = 54, -4, 8; T= 5.20).

In a second contrast, we isolated prosodic processing when the prosodic cue was superfluous for the sentence structure (*BL_ProsNoChoice* vs *ProsOnly2).* In the condition *BL_ProsNoChoice*, the task did not require processing of the intonational cue to disambiguate the sentence structure, whereas condition *ProsOnly2* was a matching sentence in which the IPB was necessary for building the sentence structure. The processing of a superfluous intonational cue showed an overall pattern of right-lateralised activity. Functional activation of the inferior frontal gyrus was right-lateralised, with peak activations in the pars opercularis (x, y, z = 51, 20, 8; T = 5.19) and pars triangularis (x, y, z = 48, 44, 8; T = 5.45) (Figure 4B). Additional right-lateralised activations were found in the superior temporal sulcus and gyrus (x, y, z = 57, -37, 8; T = 4.34), the supplementary motor area (x, y, z = 9, 23, 50; T = 7.24), and the precuneus (x, y, z = 6, -55, 41; T = 6.01). Activity in the pre/post-central gyrus (x, y, z = -39, -16, 50; T = 4.94) was stronger in the left than right hemisphere.
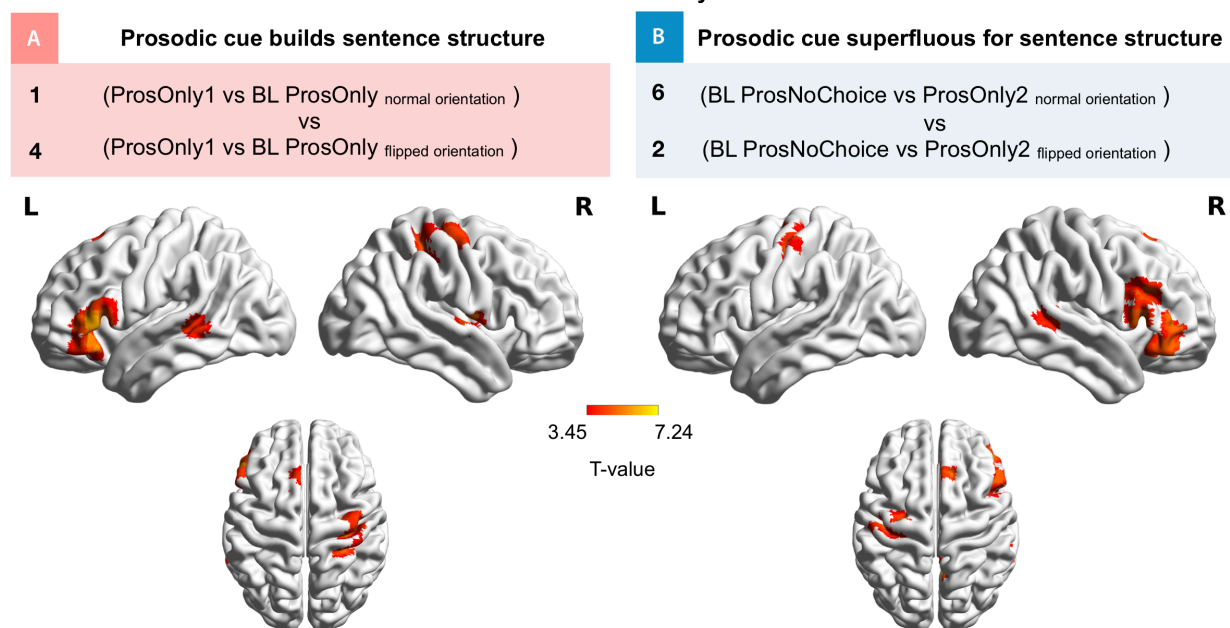
Figure 4

Lateralisation analysis showing functional contrasts of interest compared to their left-right flipped equivalent. (**A**) Lateralised functional activity evoked by processing of sentence structure guided by a prosodic cue. (**B**) Lateralised functional activity evoked by processing a sentence structure in which the prosodic cue is superfluous. All comparisons are thresholded on the cluster level at $p < 0.05$, FWE-corrected. BL stands for baseline.

Table 1 Task-related activity for the comparisons of interest thresholded on the cluster level at $p < 0.05$, FWE-corrected.

Region	hemisphere	MNI coordinates (x y z; in mm)			T	cluster size
Prosodic cue establishes sentence structure (ProsOnly1 (1) > BL_ProsOnly (4))						
Inferior frontal gyrus	L	-45	29	-4	4.93	301
Grammatical cue establishes sentence structure (GramOnly (3) > BL_GramOnly (5))						
Superior temporal sulcus	L	-57	-16	2	5.23	190
Inferior frontal gyrus	L	-51	11	8	4.43	123

Table 2 Task-related activity for the comparisons of interest in the lateralisation analysis. All results were thresholded on the cluster level at $p < 0.05$, FWE-corrected.

Region	hemisphere	MNI coordinates (x y z; in mm)			T	cluster size
Prosodic cue establishes sentence structure (ProsOnly1 (1) vs BL_ProsOnly (4))						
Inferior frontal gyrus	L	-54	29	5	6.08	262
Precentral gyrus	R	30	-19	56	5.53	169
Superior temporal gyrus	R	54	-4	8	5.20	32
Supplementary motor area	L	-6	23	50	4.87	44
Superior temporal gyrus	L	-51	-34	2	4.20	35
Prosodic cue superfluous for sentence structure (BL_ProsNoChoice (6) vs ProsOnly2 (2))						
Supplementary motor area	R	9	23	50	7.24	43
Precuneus	R	6	-55	41	6.01	61
Inferior frontal gyrus, pars triangularis	R	48	44	8	5.45	96
Inferior frontal gyrus, pars opercularis	R	51	20	8	5.19	92
Precentral gyrus	L	-39	-16	50	4.94	147
Superior temporal gyrus	R	57	-37	8	4.34	38

4. Discussion

In this study we show that the left inferior frontal gyrus has a key role in processing prosodic information that is used for sentence comprehension. By comparing the role of intonational and grammatical cues in sentence processing, we provide novel evidence that intonation is processed in the left hemisphere when its function is syntactic.

Our first aim was to investigate whether different types of language cue available to understand the structure of a sentence determined the recruitment of different brain areas. As a main finding, we show that the left IFG is involved in sentence processing both when intonational and grammatical cues establish the sentence structure. As a second finding, we show that lateralisation of activity depends on whether or not intonation is decisive for the interpretation of the sentence structure. When intonation was the only cue establishing the sentence structure, activity in the IFG was left-lateralised. Conversely, activity in the IFG was lateralised to the right hemisphere when the intonational cue was superfluous for sentence comprehension, even though the cues were acoustically identical.

Our results substantially extend previous neuroimaging work that emphasised the importance of the left IFG for sentence comprehension (Bornkessel-Schlesewsky & Schlesewsky, 2013; Friederici, 2011; Hagoort, 2014). We show that the left IFG plays a major role when an intonational cue is used to build sentence structure. Previous neuroimaging studies on prosody processing reported the left IFG as part of a wider network of bilateral fronto-temporal regions (reviewed in Belyk & Brown (2014)). In contrast, in this study we find the left IFG in isolation. This difference from previous work is likely due to two aspects of prosody which were tackled in the present study but have been largely ignored in the literature to date. First, many previous neuroimaging studies have focused on the acoustic aspect of prosody processing rather than its role in guiding sentence comprehension. A predominantly right-hemispheric temporal network has often been identified in prosody experiments drawing comparisons between normal speech and acoustically manipulated speech, such as speech with flattened pitch (Meyer et al., 2004) or filtered speech in which only the pitch contour remained (Hesling, Clément, Bordessoules, & Allard, 2005; Meyer et al., 2002; 2004). The temporal areas found in these studies are likely to reflect processing of acoustic properties of linguistic prosody, and in particular of pitch. In contrast to these previous studies, we presented an intonational cue that was acoustically identical in our condition of interest (in which the cue was used for sentence comprehension; *ProsOnly1*)

and in its matched control condition. Our finding of activity in the left IFG without additional activity in auditory regions can be explained by the acoustic similarity of these two conditions. As another novel aspect, our study investigated how intonation is used to guide the interpretation of the sentence. Notably, linguistic prosody can come in various forms (Cutler et al., 1997), of which marking of a syntactic boundary by an intonational phrase boundary (IPB) is arguably the most important for sentence structure. The importance of prosodic information in syntactic phrasing has been demonstrated in electrophysiological studies (Friederici et al., 2007; Steinhauer et al., 1999). Previous fMRI studies, however, have not studied this use of prosody. Rather, fMRI research has focused on types of linguistic prosody which are not as crucial for the syntactic structure of a sentence, such as marking a question or statement (Kreitewolf et al., 2014; Sammler et al., 2015) or placing stress (Kristensen et al., 2013; Perrone-Bertolotti et al., 2013). Moreover, studies often used a low-level or non-linguistic baseline condition rather than a comparable linguistic control task (Meyer et al., 2002; 2004; Plante et al., 2002). Right fronto-temporal areas have been found in question/statement versus phoneme discrimination tasks (Kreitewolf et al., 2014; Sammler et al., 2015), with functional activity switching to the left hemisphere when contrasted against a non-linguistic task of speaker identification (Kreitewolf et al., 2014). In turn, processing of pitch focus was shown to involve bilateral frontal and superior temporal regions (Kristensen et al., 2013; Perrone-Bertolotti et al., 2013), rather than the isolated recruitment of the left IFG found here. Critically, in the aforementioned studies, analysis of pitch differences was required to deduce linguistic meaning from the speech signal (i.e., by distinguishing question/statements or establishing constituent focus), but it was not used for the interpretation of the syntactic structure. In sum, we argue that the type of linguistic prosody in the current study (i.e., the use of the IPB) forms a more direct link to sentence structure processing than the previous studies, thus isolating functional activation in the left IFG.

Our results show that the left IFG was also involved when a grammatical cue (i.e., word form) guided sentence comprehension. In the *grammatical cue* condition *(GramOnly)*, the sentence structure could only be resolved by matching the case of the personal pronoun (morphosyntactic information) to either one of the verbs in the sentence. Although different from the intonational cue, both cues had the syntactic function of unambiguously attributing the object in the sentence to one of two verbs. Case marking ensured that the personal

pronoun could only match one of two verbs, similar to how the position of the IPB established a single possible interpretation of the sentence. We found that the left IFG was engaged in the processing of both cue types, which indicates that this area responds to different kinds of cues resolving ambiguity in sentence structure. This finding points to a more general involvement of the left IFG in the processing of sentence structure. Activity in the left STG/STS, on the other hand, was only present when sentence structure was built by a grammatical cue. This is not surprising given that superior temporal regions have been associated with morphosyntactic processing in lesions studies (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004) and fMRI studies on assigning subject and object to a verb (Bornkessel et al., 2005). Additionally, the posterior STG has been shown to play a critical role in the production of the correct morphosyntactic form (D. K. Lee et al., 2018).

To isolate the processing of each type of linguistic cue, we compared an experimental condition in which only one cue (either intonational or grammatical) was available to resolve sentence structure ambiguity, to a control condition in which both cues were available. These comparisons did not only reveal differences in the neural activity but also on the behavioural level: participants responded faster and more accurately in the control conditions compared to the experimental conditions. Consequently, one could argue that the increased activity in the left IFG reflects differences in task difficulty rather than differences in linguistic processing per se. However, our general linear model included a regressor which modelled all trial-by-trial responses as a boxcar function with trial-by-trial reaction times as duration (Grinband et al., 2008). Since one regressor was built modelling reaction times across all conditions, this regressor should account for variance introduced by between-condition differences in reaction times and should therefore regress out domain general effects. Moreover, increased cognitive demand usually relies on a domain-general network that in the frontal lobe excludes most of the IFG (Fedorenko, Duncan, & Kanwisher, 2013) and rather includes premotor regions, the anterior cingulate cortex, and the middle frontal gyrus (Duncan, 2010). Taken together, we did not find support for the alternative explanation that the observed activity in left IFG was due to differences in task difficulty.

With respect to the lateralisation of intonation processing, we set out to advance a debate that has been held for decades. Although early studies indicated a right-hemispheric advantage for emotional prosody and left-hemispheric dominance for linguistic prosody (Heilman, Bowers, Speedie, & Coslett, 1984; Luks et al., 1998), recent meta-analyses suggest

involvement of a more bilateral network (Belyk & Brown, 2014; Witteman et al., 2011). Our paradigm allowed us to assess the lateralisation of prosody processing in function of its linguistic importance since we varied the linguistic role of the intonational cue across conditions while keeping acoustical information identical across conditions. We found that processing intonation was left-lateralised in the IFG when it guided sentence comprehension. In contrast, when intonation was superfluous for disambiguation of the sentence structure, activation in the IFG was shifted to the right hemisphere. The latter finding can be explained by our manipulation in which the IPB was not relevant for disambiguating the sentence structure. This resulted in a relative dominance of the right IFG when prosody was processed without being used to establish the sentence structure. This interpretation is consistent with previous studies (Kreitewolf et al., 2014; Sammler et al., 2015), demonstrating right IFG involvement when intonational contours were processed without requiring integration into a sentence structure. The observed right-hemispheric lateralisation of the pSTS in our study further converges with previous work (Meyer et al., 2002; 2004; Sammler et al., 2015) and with models describing a right-hemispheric dominance of auditory regions in processing pitch information in speech, such as intonational contours (Poeppel, 2003). Together, the previous and present results suggest that the right pSTS is preferentially involved in processing of intonational contours as such, but not in the subsequent integration of this information during sentence comprehension.

Conclusion

In summary, our results provide evidence for a key role of the left IFG in sentence processing when only intonation conveys the structure of the sentence. Activity in this region overlapped with the region that was active when the sentence structure was established by a grammatical cue, i.e. word form. This finding extends previous work on the contribution of the left IFG in sentence comprehension, highlighting the role of this region in the integration of prosodic as well as grammatical cues into the sentence structure. Moreover, we found that lateralisation of intonation processing depends on whether or not intonation is critical for understanding a sentence structure. This supports the notion that processing of prosodic information is lateralised in function of its linguistic role (Friederici & Alter, 2004; Kreitewolf et al., 2014; Luks et al., 1998; van Lancker, 1980), showing this distinction for the first time in an fMRI study using sentence-level intonation in natural speech.

References

Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: an ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *9*(9), 1395–1403. http://doi.org/10.1093/scan/nst124

Bornkessel, I., Zysset, S., Friederici, A. D., Cramon, von, D. Y., & Schlesewsky, M. (2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, *26*(1), 221–233. http://doi.org/10.1016/j.neuroimage.2005.01.032

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Reconciling time, space and function: a new dorsal-ventral stream model of sentence comprehension. *Brain and Language*, *125*(1), 60–76. http://doi.org/10.1016/j.bandl.2013.01.010

Bozic, M., Tyler, L. K., Ives, D. T., Randall, B., & Marslen-Wilson, W. D. (2010). Bihemispheric foundations for human speech comprehension. *Proceedings of the National Academy of Sciences*, *107*(40), 17439–17444. http://doi.org/10.1073/pnas.1000531107

Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D. J. (2011). Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials. *Language and Linguistics Compass*. http://doi.org/10.1111/j.1749-818X.2011.00291.x

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech*, *40 (Pt 2)*(2), 141–201. http://doi.org/10.1177/002383099704000203

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Jr., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1-2), 145–177. http://doi.org/10.1016/j.cognition.2003.11.002

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends in Cognitive Sciences*, *14*(4), 172–179. http://doi.org/10.1016/j.tics.2010.01.004

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, *25*(4), 1325–1335. http://doi.org/10.1016/j.neuroimage.2004.12.034

Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., & Bolhuis, J. J. (2015). Structures, Not Strings: Linguistics as Part of the Cognitive Sciences. *Trends in Cognitive Sciences*, *19*(12), 729–743. http://doi.org/10.1016/j.tics.2015.09.008

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, *110*(41), 16616–16621. http://doi.org/10.1073/pnas.1315235110

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*(4), 1357–1392. http://doi.org/10.1152/physrev.00006.2011

Friederici, A. D., & Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain and Language*, *89*(2), 267–276. http://doi.org/10.1016/S0093-934X(03)00351-1

Friederici, A. D., Cramon, von, D. Y., & Kotz, S. A. (2007). Role of the corpus callosum in speech comprehension: interfacing syntax and prosody. *Neuron*, *53*(1), 135–145. http://doi.org/10.1016/j.neuron.2006.11.020

Friston, K. J., Worsley, K. J., Frackowiak, R. S., Mazziotta, J. C., & Evans, A. C. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping*, *1*(3), 210–220. http://doi.org/10.1002/hbm.460010306

Goucha, T., & Friederici, A. D. (2015). The language skeleton after dissecting meaning: A functional segregation within Broca's Area. *NeuroImage*, *114*, 294–302. http://doi.org/10.1016/j.neuroimage.2015.04.011

Grinband, J., Wager, T. D., Lindquist, M., Ferrera, V. P. & Hirsch, J, (2008). Detection of time-varying signals in event-related fMRI designs. *NeuroImage*, 43, 509–520.

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology*, *28*, 136–141. http://doi.org/10.1016/j.conb.2014.07.013

Heilman, K. M., Bowers, D., Speedie, L., & Coslett, H. B. (1984). Comprehension of affective and nonaffective prosody. *Neurology*, *34*(7), 917–917. http://doi.org/10.1212/WNL.34.7.917

Hesling, I., Clément, S., Bordessoules, M., & Allard, M. (2005). Cerebral mechanisms of prosodic integration: evidence from connected speech. *NeuroImage*, *24*(4), 937–947. http://doi.org/10.1016/j.neuroimage.2004.11.003

Ischebeck, A. K., Friederici, A. D., & Alter, K. (2008). Processing Prosodic Boundaries in Natural and Hummed Speech: An fMRI Study. *Cerebral Cortex (New York, N.Y. : 1991)*, *18*(3), 541–552. http://doi.org/10.1093/cercor/bhm083

Josse, G., Kherif, F., Flandin, G., Seghier, M. L., & Price, C. J. (2009). Predicting language lateralization from gray matter. *Journal of Neuroscience*, *29*(43), 13516–13523. http://doi.org/10.1523/JNEUROSCI.1680-09.2009

Kreitewolf, J., Friederici, A. D., & Kriegstein, von, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2014.07.038

Kristensen, L. B., Wang, L., Petersson, K. M., & Hagoort, P. (2013). The interface between language and attention: prosodic focus marking recruits a general attention network in spoken language comprehension. *Cerebral Cortex*, *23*(8), 1836–1848. http://doi.org/10.1093/cercor/bhs164

Lee, D. K., Fedorenko, E., Simon, M. V., Curry, W. T., Nahed, B. V., Cahill, D. P., & Williams, Z. M. (2018). Neural encoding and production of functional morphemes in the posterior temporal lobe. *Nature Communications*, *9*(1), 1877. http://doi.org/10.1038/s41467-018-04235-3

Liégeois, F., Connelly, A., Salmond, C. H., Gadian, D. G., Vargha-Khadem, F., & Baldeweg, T. (2002). A direct test for lateralization of language activation using fMRI: comparison with

invasive assessments in children with epilepsy. *NeuroImage*, *17*(4), 1861–1867. http://doi.org/10.1006/nimg.2002.1327

Luks, T. L., Nusbaum, H. C., & Levy, J. (1998). Hemispheric involvement in the perception of syntactic prosody is dynamically dependent on task demands. *Brain and Language*, *65*(2), 313–332. http://doi.org/10.1006/brln.1998.1993

Macmillan, N. A. & Creelman, C. D. Detection Theory. (Cambridge University Press, 1991).

Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic Effects in Minimal Attachment. *The Quarterly Journal of Experimental Psychology. a, Human Experimental Psychology*, *45*(1), 73–87. http://doi.org/10.1080/14640749208401316

Männel, C., & Friederici, A. D. (2009). Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *Journal of Cognitive Neuroscience*, *21*(10), 1988–2006. http://doi.org/10.1162/jocn.2009.21221

Meyer, L., Henry, M. J., Gaston, P., Schmuck, N. & Friederici, A. D (2016). Linguistic Bias Modulates Interpretation of Speech via Neural Delta-Band Oscillations. *Cerebral Cortex*. doi:10.1093/cercor/bhw228

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & Cramon, von, D. Y. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, *17*(2), 73–88. http://doi.org/10.1002/hbm.10042

Meyer, M., Steinhauer, K., Alter, K., Friederici, A. D., & Cramon, von, D. Y. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain and Language*, *89*(2), 277–289. http://doi.org/10.1016/S0093-934X(03)00350-X

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Perrone-Bertolotti, M., Dohen, M., Lœvenbruck, H., Sato, M., Pichat, C., & Baciu, M. (2013). Neural correlates of the perception of contrastive prosodic focus in French: A functional magnetic resonance imaging study. *Human Brain Mapping*, *34*(10), 2574–2591. http://doi.org/10.1002/hbm.22090

Plante, E., Creusere, M., & Sabin, C. (2002). Dissociating sentential prosody from sentence processing: activation interacts with task demands. *NeuroImage*, *17*(1), 401–410. http://doi.org/10.1006/nimg.2002.1182

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time." *Speech Communication*, *41*(1), 245–255. http://doi.org/10.1016/S0167-6393(02)00107-3

Regel, S., Kotz, S. A., Henseler, I., & Friederici, A. D. (2017). Left inferior frontal gyrus mediates morphosyntax: ERP evidence from verb processing in left-hemisphere damaged patients. *Cortex*, *86*, 156–171. http://doi.org/10.1016/j.cortex.2016.11.007

Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. *Journal of Neuroscience*, *31*(47), 16949–16957. http://doi.org/10.1523/JNEUROSCI.2725-11.2011

Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, P. E. G., & Belin, P. (2015). Dorsal and Ventral Pathways for Prosody. *Current Biology : CB*, *25*(23), 3079–3085. http://doi.org/10.1016/j.cub.2015.10.009

Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V., & Friederici, A. D. (2010). Prosody meets syntax: the role of the corpus callosum. *Brain*, *133*(9), 2643–2655. http://doi.org/10.1093/brain/awq231

Selkirk, E. O. (1984). Phonology and syntax. MIT Press.

Sportiche, D., Koopman, H., & Stabler, E. (2013). An Introduction to Syntactic Analysis and Theory. John Wiley & Sons.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*(2), 191–196. http://doi.org/10.1038/5757

Strelnikov, K. N., Vorobyev, V. A., Chernigovskaya, T. V., & Medvedev, S. V. (2006). Prosodic clues to syntactic processing—a PET and ERP study. *NeuroImage*, *29*(4), 1127–1134. http://doi.org/10.1016/j.neuroimage.2005.08.021

Tang, C., Hamilton, L. S., & Chang, E. F. (2017). Intonational speech prosody encoding in the human auditory cortex. *Science*, *357*(6353), 797–801. http://doi.org/10.1126/science.aam8577

van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social Interaction*. http://doi.org/10.1080/08351818009370498

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432. http://doi.org/10.1016/j.neuroimage.2005.11.002

Wildgruber, D., Hertrich, I., Riecker, A., Erb, M., Anders, S., Grodd, W., & Ackermann, H. (2004). Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. *Cerebral Cortex (New York, N.Y. : 1991)*, *14*(12), 1384–1389. http://doi.org/10.1093/cercor/bhh099

Witteman, J., van IJzendoorn, M. H., van de Velde, D., van Heuven, V. J. J. P., & Schiller, N. O. (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: A meta-analysis of the lesion literature. *Neuropsychologia*, *49*(13), 3722–3738. http://doi.org/10.1016/j.neuropsychologia.2011.09.028

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PloS One*, *8*(7), e68910. http://doi.org/10.1371/journal.pone.0068910

Zatorre, R. J. (2001). Neural specializations for tonal processing. *Annals of the New York Academy of Sciences*, *930*, 193–210.

Supplementary materials Van der Burght *et al.*
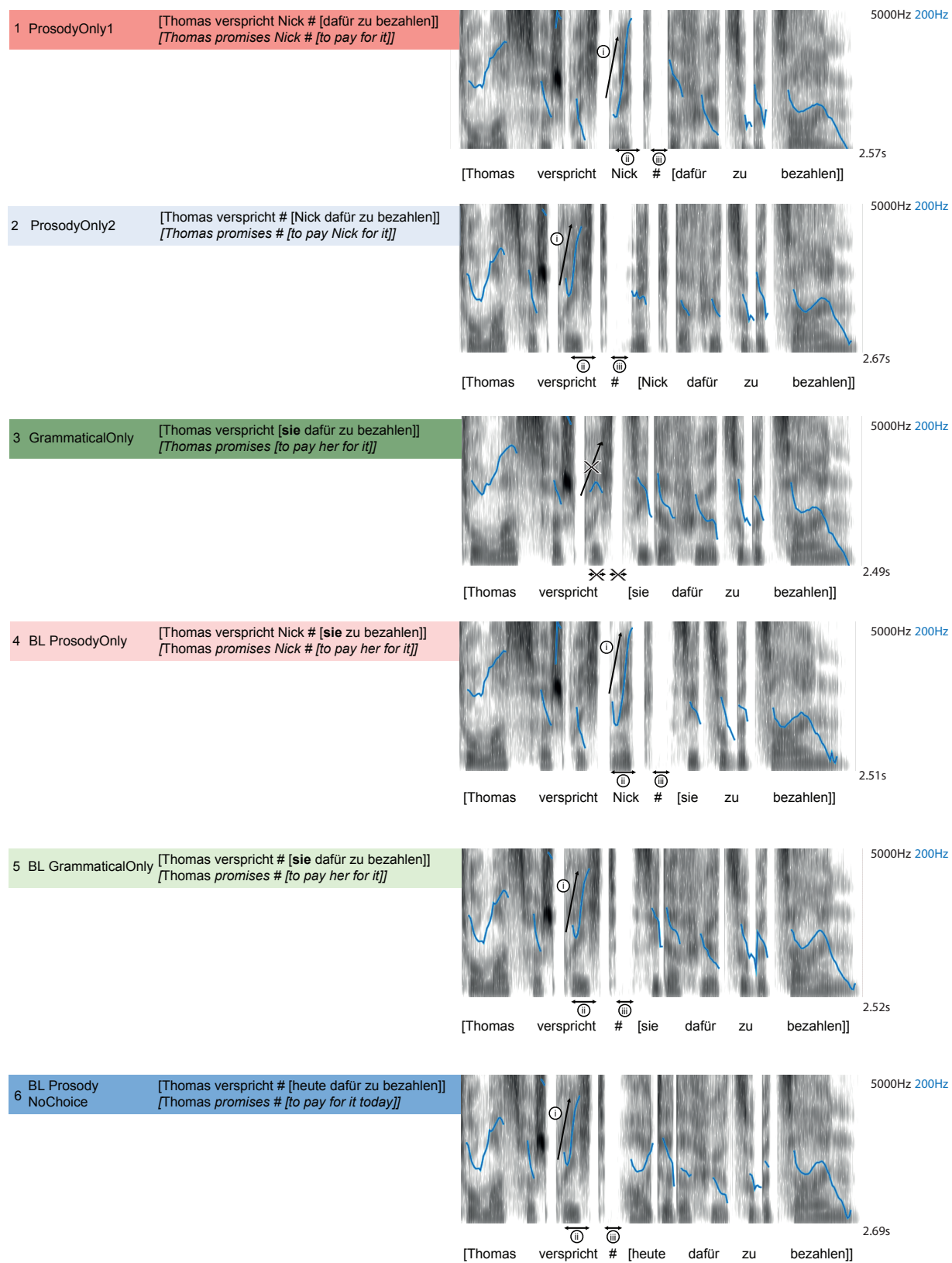*Intonation guides sentence processing in the left inferior frontal gyrus*

Supplementary table 1 Reaction times. Pair-wise comparisons (Wilcoxon signed-rank tests) between reaction times in all experimental conditions. Table shows *Z*-scores followed by the *P*-value per comparison. *P*-values smaller than 0.0033 (Bonferroni-corrected) are marked with an †.

	ProsSyn1	ProsSyn2	GramOnly	BL_Pros Only	BL_Gram Only	BL_Pros_ NoChoice
ProsSyn1	-	-	-	-	-	-
ProsSyn2	-0.597; 0.551	-	-	-	-	-
GramOnly	-1.283; 0.200	-2.121; 0.034	-	-	-	-
BL_ProsOnly	-4.457; <0.001†	-4.457; <0.001†	-4.432; <0.001†	-	-	-
BL_GramOnly	-3.594; <0.001†	-3.060; 0.002†	-4.026; <0.001†	-4.432; <0.001†	-	-
BL_Pros_ NoChoice	-4.203; <0.001†	-4.178; <0.001†	-4.203; <0.001†	-2.197; 0.028	-3.822; <0.001†	-

Supplementary table 2 Accuracy rates. Pair-wise comparisons (Wilcoxon signed-rank tests) between accuracy rates in all experimental conditions. Table shows *Z*-scores followed by the *P*-value per comparison. *P*-values smaller than 0.0033 (Bonferroni-corrected) are marked with †.

	ProsSyn1	ProsSyn2	GramOnly	BL_Pros Only	BL_Gram Only	BL_Pros_ NoChoice
ProsSyn1	-	-	-	-	-	-
ProsSyn2	-1.939; 0.053	-	-	-	-	-
GramOnly	-0.633; 0.527	-2.331; 0.020	-	-	-	-
BL_ProsOnly	-3.523; <0.001†	-2.578; 0.010	-3.526; <0.001†	-	-	-
BL_GramOnly	-3.023; 0.003†	-1.882; 0.259	-3.760; <0.001†	-1.882; 0.060	-	-
BL_Pros_ NoChoice	-4.270; <0.001†	-3.816; <0.001†	-3.936; <0.001†	-2.024; 0.043	-2.908; 0.004	-

Supplementary Figure 1

1 ProsodyOnly1
[Thomas verspricht Nick # [dafür zu bezahlen]]
[Thomas promises Nick # [to pay for it]]

5000Hz 200Hz

2.57s

[Thomas] [verspricht] [Nick] [#] [dafür] [zu] [bezahlen]]

2 ProsodyOnly2
[Thomas verspricht # [Nick dafür zu bezahlen]]
[Thomas promises # [to pay Nick for it]]

5000Hz 200Hz

2.67s

[Thomas] [verspricht] [#] [Nick] [dafür] [zu] [bezahlen]]

3 GrammaticalOnly
[Thomas verspricht [**sie** dafür zu bezahlen]]
[Thomas promises [to pay her for it]]

5000Hz 200Hz

2.49s

[Thomas] [verspricht] [[sie] [dafür] [zu] [bezahlen]]

4 BL ProsodyOnly
[Thomas verspricht Nick # [**sie** zu bezahlen]]
[Thomas promises Nick # [to pay her for it]]

5000Hz 200Hz

2.51s

[Thomas] [verspricht] [Nick] [#] [sie] [zu] [bezahlen]]

5 BL GrammaticalOnly
[Thomas verspricht # [**sie** dafür zu bezahlen]]
[Thomas promises # [to pay her for it]]

5000Hz 200Hz

2.52s

[Thomas] [verspricht] [#] [sie] [dafür] [zu] [bezahlen]]

6 BL Prosody NoChoice
[Thomas verspricht # [heute dafür zu bezahlen]]
[Thomas promises # [to pay for it today]]

5000Hz 200Hz

2.69s

[Thomas] [verspricht] [#] [heute] [dafür] [zu] [bezahlen]]

Caption Supplementary Figure 1
Spectrograms with pitch contours for each of the experimental conditions. In the five conditions with intonational phrase boundary (IPB), three acoustic events can be observed: (i) a pitch rise and (ii) syllable lengthening, followed by (iii) a pause. Note that in the Grammatical Only condition the IPB is absent, marked by an absence of pitch rise, syllable lengthening and pause (indicated with an X).