ON THE ACOUSTICS OF *WH*-EXCLAMATIVES AND *WH*-INTERROGATIVES: EFFECTS OF INFORMATION STRUCTURE AND SEX OF SPEAKER[1]

Sophie Repp

Humboldt-Universität zu Berlin
sophie.repp@rz.hu-berlin.de

ABSTRACT

In German, *wh*-questions and verb-second *wh*-exclamatives are string-identical and can only be distinguished by intonation. This study presents data from a production experiment where speakers produced contextualized questions and exclamatives, showing that the two sentence types differ in many acoustic measures throughout the clause. The results also indicate that the realization of both sentence types is context-dependent in terms of information structure: focus is marked both in *wh*-questions and *wh*-exclamatives. Finally, the study shows that female speakers mark the difference between exclamativity and interrogativity more strongly than male speakers do.

Keywords: intonation, exclamative, question, information structure, focus, givenness, sex of speaker

1. INTRODUCTION

In German, like in other West-Germanic languages, the intonation of an utterance is mainly determined by its information structure (IS) and by its illocutionary force, i.e. whether it is an assertion, a question or e.g. an expressive speech act. Both IS and certain speech act types (assertions, questions) have been studied extensively for their prosodic effects but the interaction of these factors is largely unexplored: IS has been studied mainly in assertions.

The current study investigates focus and givenness in verb-second *wh*-clauses in German, which can be interpreted as exclamatives or questions depending on their intonation. In assertions, new information focus is usually marked by a H* accent ([3], [10], [24]), which is raised for narrow vs. broad focus, and the accented syllable is lengthened ([5], [11]). Given expressions are typically deaccented, or if prenuclear, marked with a L* accent ([4], [20]).

Exclamatives typically come with a so-called *exclamative accent (EA)*, which in *wh*-exclamatives has been claimed to occur on the scalar expression that often occurs in these sentences ([1], cf. English *How TALL she is!*). [1] furthermore proposes that another, potentially focussing accent may occur on a 'suitable' element towards the end of the clause. If there is no scalar expression, the verb in verb-second position carries the EA. It is worth noting that for non-*wh*-exclamatives the EA is often assumed to occur on a strong (e.g. subject) *d*-pronoun (e.g. [1], see below). As for the acoustics of the EA, [2] reports for exclamatives in declarative sentence form that a later pitch peak, and a combination of higher duration and higher pitch peak help listeners to identify the accent as an EA rather than a focus accent. Whether or not the EA is related to focus is debated. Both its location [17], and its phonetic characteristics [2] have been suggested not to be influenced by IS. However, [9] argues that in polar exclamatives the position of the EA varies with the focus-licensing context. Generally, exclamatives seem to be longer than assertions [1]. They end in a fall.

The role of focus in German *wh*-questions has received quite some attention in the semantic literature (e.g. [4][6], [13], [18], [23]). The *wh*-word is often assumed to be focussed ([6], [13]), but this focus has been claimed not to be realized by an accent ([1], [13]) – unless the remainder of the clause is contextually given, i.e. has a specific IS ([18], [23]). Quantitative empirical investigations of Dutch ([8], [12]) and English *wh*-questions ([14]) provide initial evidence that the IS of the clause indeed plays a role for the prosodic realization of the *wh*-word, just as it does for the realization of the whole clause. For German these assumptions have not been tested quantitatively. *Wh*-questions may end with a fall or a rise.

The present study explores the intonational characteristics of *wh*-exclamatives and *wh*-questions in interaction with focus by examining the acoustics of the two sentence types in a controlled production experiment where speakers uttered target sentences in an appropriate linguistic context. It examines what prosodic means speakers employ to differentiate questions and exclamatives, specifically, if the *wh*-pronoun, the auxiliary or the *d*-pronoun in the clause-initial region are realized differently in the two sentence types. Furthermore, it tests if and how focus/givenness is marked in the two sentence types. The study was carried out with male and female participants to explore potential effects of sex of speaker, which has been reported to influence e.g. the

marking of interrogativity in Dutch ([25]) as well as focus marking in German ([20], [22]).

2. METHOD

2.1. Design and materials

The study had a 2×2×2 mixed design with the factors SENTENCE TYPE (ST, levels: *exclamative, interrogative*), IS (*focussed object, non-focussed object*), and SEX (*male, female*), yielding four within-subjects conditions, eight conditions altogether. The target sentence (see below for illustration) was a verb-second clause with an adverbial *wh*-pronoun in clause-initial position, followed by an auxiliary in second position. The sentence also contained a subject, an object, a main verb in clause-final position and some adverbs. The subject was a *d*-pronoun, which in non-*wh* exclamatives often carries the EA, and which in oral German is used as a strong personal pronoun irrespective of sentence type.

Wo	hat	die	schon	überall	Aromen	entdeckt
where	has	she	already	everywhere	flavourings	found
wh	*aux*	*d-pron*	*--- adv---*		*object*	*verb*

The discourse context, consisting of a dialogue between two speakers (see below), licensed either an exclamative or an interrogative interpretation of the target sentence (which was also marked by an exclamation mark / question mark). In the conditions where the object in the target sentence (e.g. *flavourings*) was not a focus, the context explicitly introduced that object referent, i.e. made it given. In the conditions with a focussed object, the context introduced an implicit set of referents of which the object referent was a plausible member (e.g. *food additives*): in this type of context the object in the target sentence is a focus because the context provides an antecedent for the focus alternatives ([21]).

Speaker A (exclamative & interrogative conditions): Julia kontrolliert Lebensmittel und findet oft merkwürdige Dinge in vielerlei Lebensmitteln. Sie ist eine Spezialistin beim Nachweisen der verschiedensten {Zusatzstoffe$_{\text{FOCUS}}$/Aromen$_{\text{NON-FOCUS}}$.}
'Julia works as a food safety manager and often finds strange things in all sorts of foods. She is a specialist when it comes to detecting sundry {additives$_{\text{FOCUS}}$ / flavourings$_{\text{NON-FOCUS}}$}'.

Speaker B (exclamative conditions): Ja, Julia ist sehr gewissenhaft. Sie sucht unheimlich gründlich. *'Yes, Julia is very conscientious. She searches very rigorously.'*
Wo hat die schon überall Aromen entdeckt!
'The foods that she's already found flavourings in!'

Speaker B (interrogative conditions): Im Ernst? Da muss Julia sicherlich gründlich suchen. Was weißt du denn darüber? *('Really? Julia must search very rigorously. What do you know about this?')*
Wo hat die schon überall Aromen entdeckt?
'Where has she already found flavourings?'

The test materials consisted of eight lexicalizations with the same metrical structure as the above example, resulting in 32 experimental items, as well as 16 filler items from another experiment.

2.2. Participants and procedure

16 speakers of Standard German from the Berlin-Brandenburg area (8 male, 8 female, mean age 23.6) took part in the experiment. The experiment was run using the software *Presentation* (Neurobehavioral systems). Participants took the role of the second speaker in the dialogues. First, they heard and read the pre-recorded text of the first speaker. Then they quietly read the reply of the second speaker. When they felt they had understood the reply they recorded it. They were asked to speak in the way that they found most natural in the given context. It was pointed out to them that some utterances might be rather passionate (*emotional* in German). Items were recorded in a pseudo-randomized order.

2.3. Analysis

19 of 512 utterances (3.7 % data points) were discarded due to hesitations, technical errors or speech errors. The 493 utterances left for analysis were annotated by hand for syllable and word boundaries in PRAAT [7]. Voice pulses were corrected manually. Acoustic measures were drawn from the data using the PRAAT script ProsodyPro [26], which was also used to create the F0 graph in Figure 1.

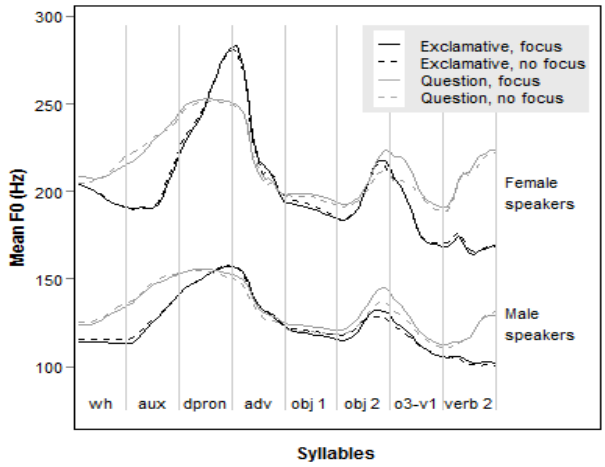
For the acoustic investigation, maximum, minimum and mean F0 ($F0_{max}$, $F0_{min}$, $F0_{mean}$), F0 excursion ($F0_{exc}$), duration and intensity were taken for the following utterance segments: *wh*-pronoun, auxiliary, *d*-pronoun, first and second (= stressed) syllable of the object, the second (= stressed) syllable of the main verb, and the stretch between the two stressed syllables, comprising the third syllable of the object (which was not always realized) and the first syllable of the verb. Furthermore, pitch peak position was determined for the region auxiliary–*d*-pronoun, as well as for the object region. The statistical analysis was carried out on normalized data: for each segment, linear regression models with speaker and lexicalization as predictors were computed, and by subtracting the value predicted by the models from the actual value in each production residual values were obtained. These residual values were subjected

to statistical analysis (multilevel linear models with the factors ST and IS nested within participants; R package nmle [19]). For the clause-initial region (*wh*-pronoun, auxiliary, *d*-pronoun), an additional auditory investigation was conducted to determine the accent location. For the frequency distribution a four-way log-linear analysis was conducted.

3. RESULTS AND DISCUSSION

Figure 1 shows the time-normalized F0 contour for the four different sentence types investigated, separate for male and female speakers. Figure 2 shows the mean residual $F0_{exc}$: IS mainly manifested itself in $F0_{exc}$ (and $F0_{max}$), see below.

Figure 1: Time-normalized F0 contour averaged



across female vs. male speakers.

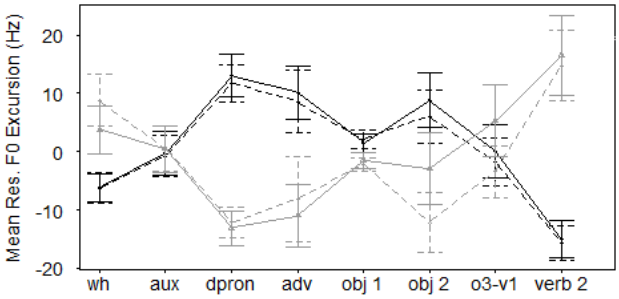
The clause-initial region

Statistical analysis for the ***wh*-pronoun** revealed a main effect of ST for all measures that were taken (duration: $\chi^2(6) = 6.6$, $p = .01$, $F0_{max}$/ $F0_{min}$/ $F0_{mean}$/ $F0_{exc}$/ duration/ intensity: at least: $\chi^2(6) > 11$, $p < .001$). The values for all measures where higher in questions than in exclamatives. For $F0_{max}$ and duration there was a marginal main effect of IS ($F0_{max}$: $\chi^2(7) = 3.1$, $p = .078$; duration: $\chi^2(7) = 3.18$, $p = .07$): $F0_{max}$ was higher and duration was longer if the object referent was non-focus rather than focus. $F0_{exc}$ showed a marginal interaction ST×IS ($\chi^2(8) = 5.56$, $p = .06$). In questions, $F0_{exc}$ on the *wh*-pronoun was marginally higher if the object was non-focus than when it was focus ($b = 2.6$, $t(15) = 1.82$, $p = .089$). In exclamatives, IS had no effect on $F0_{exc}$.

For the **auxiliary**, statistical analysis revealed a main effect of ST for all measures except $F0_{exc}$ (at least $\chi^2(6) > 12$, $p < .001$). The values of these measures were higher in questions than in exclamatives. For $F0_{min}$ there was a marginal effect of IS ($\chi^2(7) = 3.0$, $p = 0.08$): If the object was non-focus $F0_{min}$ on the auxiliary was higher.

For the ***d*-pronoun**, there was a main effect of ST for all measures (at least $\chi^2(6) > 4.0$, $p < .05$). $F0_{min}$ was lower in exclamatives than in questions. For $F0_{max}$, $F0_{exc}$, duration and intensity there was an interaction of ST×SEX (at least $\chi^2(8) > 7.0$, $p < .05$).[2] The values for $F0_{exc}$, duration and intensity were larger in exclamatives than in questions and the differences were greater for female than for male speakers (e.g. $F0_{exc}$: $b_{female} = 20.1$, $t(15) = 16.4$, $p < .0001$, $b_{male} = 4.8$, $t(15) = 3.1$, $p < .05$): female speakers had lower measures than male speakers in questions, and higher measures in exclamatives. $F0_{max}$ was higher in exclamatives than in questions only for female speakers ($b_{female} = 14.5$, $t(7) = 4.1$, $p < .01$, $p_{male} = .4$). The F0 peak position in the region auxiliary–*d*-pronoun was reached later in exclamatives ($\chi^2(6) = 17.3$, $p < .0001$), and by female speakers ($\chi^2(6) = 4.1$, $p < .05$).

Figure 2: Mean residual pitch excursion (95% CI)



The results indicate that speakers reliably distinguish questions from exclamatives at the beginning of the clause. The *wh*-pronoun and the auxiliary are more prominent (higher F0, longer duration, larger intensity) in questions, and the *d*-pronoun is more prominent in exclamatives, where the latter difference is significantly larger for female than for male speakers. There are indications – albeit weak ones – that the IS of the clause influences the acoustic realization esp. of the *wh*-pronoun. In line with earlier findings for *wh*-questions in Dutch and English, the *wh*-pronoun is realized more prominently if the object, which occurs later in the clause, is given rather than focussed. Since the remainder of the clause also is given, the *wh*-pronoun due to its inherent semantic focus is a good candidate for prosodic prominence. The weak interaction of ST and IS for $F0_{exc}$ suggests that the *wh*-pronoun is more susceptible to the IS of the clause in questions than in exclamatives.

The four-way log-linear analysis carried out to analyze the frequency of the accent location in the clause-initial region in relation to ST, IS and SEX (see Table 1) revealed that the effect of IS is not reflected in the accent location: there was an interaction ST×SEX×Accent location ($\chi^2(9) = 3.27$, $p < 0.05$)[3] and no effect of IS. This analysis revealed further differences between female and male speakers.

Female speakers placed the accent in exclamatives quite consistently on the *d*-pronoun whereas male speakers placed it equally often on the auxiliary, which is something both female and male speakers did in questions. Thus, contrary to descriptions in the literature, in the present study it was the *d*-pronoun rather than the verb in second position that was accented in exclamatives, at least for female speakers. Male speakers used both strategies. Note, however, that the accent distribution does not fully explain the sex differences in the acoustic measures. Even for exclamatives with the accent on the *d*-pronoun (n = 175, n_{female}: 112, n_{male} = 63), $F0_{exc}$ showed a reliable effect of SEX ($\chi^2(6) = 6.28$, $p < .01$), and $F0_{max}$ a marginal one ($\chi^2(6) = 3.16$, $p = .076$). Both measures were larger for female speakers. Finally, the ST-related findings for the acoustics of the *wh*-pronoun are reflected in the accent distribution: in questions the accent goes significantly more often to the *wh*-pronoun than in exclamatives.

Table 1: Accent location in the clause-initial region in percent for each sentence type

Accent location	**Female**		**Male**		Total
	Non-focus	Focus	Non-focus	Focus	
Exclamatives					
wh-pronoun	3.1	1.6	3.3	1.6	2.4
auxiliary	10.9	7.9	45.0	46.8	27.3
d-pronoun	85.9	90.5	51.7	51.6	70.3
Questions					
wh-pronoun	30.6	32.8	49.1	43.3	38.9
auxiliary	48.4	55.7	45.9	50.0	50.0
d-pronoun	21.0	11.5	4.9	6.7	11.1

The clause-final region: object and verb

On the **first syllable of the object**, there was a main effect of ST on $F0_{exc}$, duration, (both $\chi^2(6) > 19$, $p < .0001$) and, marginally on $F0_{min}$ ($\chi^2(6) = 3.2$, $p = .07$). Duration and $F0_{exc}$ were larger in exclamatives than in questions, $F0_{min}$ was lower.

For the **second, stressed syllable of the object**, there was a main effect of ST on $F0_{exc}$ ($\chi^2(6) = 6.6$, $p < .05$), $F0_{min}$, duration (both $\chi^2(6) > 16$, $p < .0001$), and intensity ($\chi^2(6) = 12.4$, $p < .001$). Exclamatives had a larger $F0_{exc}$ and intensity and a longer duration than questions, and a lower $F0_{min}$. There also was a main effect of IS on $F0_{max}$, $F0_{mean}$ (both $\chi^2(6) > 4$, $p < .05$), and marginally on $F0_{exc}$ ($\chi^2(6) = 3.6$, $p = .06$). The values were larger for focussed objects.

For the **stretch between the two stressed syllables**, there was a main effect of ST on $F0_{max}$, $F0_{min}$, $F0_{mean}$ and duration (all $\chi^2(6) > 8$, $p < .01$), and an interaction ST×SEX for $F0_{min}$ ($\chi^2(8) = 6.4$, $p < .05$). $F0_{max}$ and $F0_{mean}$ were lower in exclamatives than in questions, duration was longer. $F0_{min}$ was lower in exclamatives only for female speakers. There was a main effect of IS on $F0_{max}$ and $F0_{exc}$ (both $\chi^2(7) > 4.1$, $p < .05$), and a marginal interaction ST×IS for $F0_{max}$ ($\chi^2(8) = 2.9$, $p = .086$). $F0_{exc}$ was larger for focussed objects. $F0_{max}$ was higher for focussed objects only in questions (b = -5.3, $t(15)$ = -2.29, $p < .05$). Analysis of the F0 peak position in the object region only yielded marginal effects.

On the **second, stressed syllable of the main verb**, there was a main effect of ST on all measures (at least $\chi^2(6) > 12$, $p < .001$). The F0 measures as well as intensity were lower in exclamatives, duration was longer. The F0 measures also showed an interaction ST×SEX. The difference between exclamatives and questions either was only present for female speakers ($F0_{min}$) or was larger for female than for male speakers ($F0_{max}$, $F0_{mean}$, $F0_{exc}$) with female speakers having lower values for exclamatives than male speakers and higher values for questions.

The results show that speakers clearly distinguish between questions and exclamatives in the clause-final region too. On the one hand, the object is more prominent in exclamatives than in questions by several measures. Thus, exclamativity is not only signalled by the EA (e.g. on the *d*-pronoun) but also by other accent-bearing elements in the clause. On the other hand, questions end with a higher pitch and intensity but a shorter duration than exclamatives. In fact, duration is longer for exclamatives throughout the clause, except for the clause-initial *wh*-pronoun. This is compatible with previous findings according to which exclamatives are longer than assertions [1]. The results for IS suggest that focussed objects are more prominent than non-focussed objects (esp. higher F0) irrespective of sentence type. However, there also was some indication that IS is marked more clearly in questions than in exclamatives, which might be related to the overall higher prominence of the object in exclamatives which might temper with givenness marking to some extent.

4. SUMMARY

Wh-exclamatives and *wh*-questions display different acoustics throughout the clause. Female speakers mark these differences more clearly than male speakers do. The EA in *wh*-exclamatives mainly occurs on the *d*-pronoun, but also – for male speakers – on the auxiliary. IS is marked on the focussed/given element in both sentence types, but more clearly in questions, which might have to do with an additional marking of exclamativity on that element. Questions mark focus/giveness on the object also clause-initially on the *wh*-pronoun. The characteristics of the EA do not seem to interact with IS.

7. REFERENCES

[1] Altmann, H. 1993. Satzmodus. In: Jacobs, J., Stechow, A., von Sternefeld, W., Vennemann, T. (eds.), *Syntax. Ein internationales Handbuch zeitgenössischer Forschung.* Berlin: de Gruyter, 1006–1029.

[2] Batliner, A. 1988: Der Exklamativ: Mehr als Aussage oder doch nur mehr oder weniger Aussage? Experimente zur Rolle von Höhe und Position des F0-Gipfels. In: Altmann, H. (ed.), *Intonationsforschungen.* Tübingen: Niemeyer, 243–271.

[3] Batliner, Anton (1989) Fokus, Modus, und die große Zahl. Zur intonatorischen Induzierung des Fokus im Deutschen. In: Altmann, H., Batliner, A. Oppenrieder, W. (eds.), *Zur Intonation von Modus und Fokus im Deutschen*. Tübingen: Niemeyer, 21-7.

[4] Baumann, S. 2006. *The Intonation of Givenness – Evidence from German.* Tübingen: Niemeyer.

[5] Baumann, S, Grice, M., Steindamm, S. 2006. Prosodic marking of focus domains - Categorical or gradient? *Proc. SpeechProsody.* Dresden, 301-304.

[6] Beck, S. 2006.Intervention effects follow from focus interpretation. *Natural Language Semantics,* 14, 1-56.

[7] Boersma, P. Weenink, D. 2014. http://www.fon.hum.uva.nl/praat/

[8] Chen, A. 2012. Shaping the intonation of *wh*-questions: information structure and beyond. In: de Ruiter, J.P. (ed.), *Questions: Formal, Functional and Interactional Perspectives.* Cambridge: CUP, 146-162.

[9] D'Avis, F.-J. 2012. On the placement of the main accent in exclamatives. Talk presented at *Information Structure in Non-Assertive Speech Acts. Workshop at the 34th Annual Meeting of the German Linguistic Society* (DGfS), Frankfurt.

[10] Féry, C. 1993. *German Intonational Patterns.* Tübingen: Niemeyer.

[11] Féry, C. & Kügler, F. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics,* 36, 680-703.

[12] Haan. J. 2002. *Speaking of Questions. An Exploration of Dutch Question Intonation.* Utrecht: LOT Publications.

[13] Haida, A. 2007.*The Indefiniteness and Focusing of* Wh-*Words.* PhD thesis, Humboldt-Universität zu Berlin.

[14] Hedberg, N., Sosa, J. M. 2002. The prosody of questions in natural discourse. *Proc. Speech Prosody*. Illinois.

[15] Henton, C. 1989. Fact and fiction in the description of female and male pitch. *Language and Communication*, 9(4), 299-311.

[16] Hermes, D. J., Van Gestel, J. C. 1991. The frequency scale of speech intonation. *The Journal of the Acoustical Society of America*, 90(1), 97-102.

[17] Jacobs, J. 1988 Fokus-Hintergrund-Gliederung und Grammatik. In: Altmann, H. (ed.) *Intonationsforschungen.* Tübingen: Niemeyer, 89–134.

[18] Jacobs, J. 1991 Implikaturen und "alte Information" in *w*-Fragen. In: Reis, M., Rosengren, I. (eds.), *Fragesätze und Fragen.* Tübingen: Niemeyer, 201-222.

[19] Pinheiro, J., Bates, D., DebRoy S., Sarkar, D., R Core Team 2013. *nlme: Linear and Nonlinear Mixed Effects Models.* R package version 3.1-108.

[20] Röhr, C., Baumann, S. 2010. Prosodic marking of information status in German. *Proc. Speech Prosody*, Chicago.

[21] Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75-116.

[22] Schmid, C., Moosmüller, S. 2013. Gender differences in the phonetic realization of semantic focus. *Proceedings of the Prosody-Discourse Interface Conference IDP*, 119-123.

[23] Truckenbrodt, H. 2012. On the prosody of German *wh*-questions. In: Elordieta, G., Prieto, P. (eds.), *Prosody and Meaning.* Berlin: de Gruyter, 73-118.

[24] Uhmann, S. 1991. *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nichtlinearen Phonologie*. Tübingen: Niemeyer.

[25] Van Heuven, V., Haan, J. 2000. Phonetic correlates of statement versus question intonation in Dutch. In: Botinis, A. (ed.), *Intonation. Analysis, Modelling and Technology*. Dordrecht: Kluwer, 119-143.

[26] Xu, Y. 2013. ProsodyPro — A tool for large-scale systematic prosody analysis. *Proc. Tools and Resources for the Analysis of Speech Prosody (TRASP),* Aix-en-Provence.

[1] This work was supported by the German Research Foundation DFG as part of the Collaborative Research Centre (SFB) 632 *Information Structure* within project A2 at the Humboldt-Universität zu Berlin. Special thanks go to Lena Rosin and Margit Scheibel for helping with the preparation of the written materials, to Andreas Haida for speaking the text of the first speaker, to Carsten Schliewe for help with the experimental software, and to Lena Rosin, Kristina Ruthe and Margit Scheibel for help with running the experiment as well as annotating the audio files in PRAAT.

[2] The sex differences for pitch excursion remain significant when the analysis is run on raw data in the ERB frequency scale (Hermes & van Gestel's 1991 formula), or the ST scale (as in Henton 1989).

[3] The likelihood ratio of the corresponding model was $\chi^2(7) = 3.27$, $p = 0.86$.