On the intonation of Swedish rejections and rejecting questions

Heiko Seeliger, Sophie Repp

Abstract A production study of Swedish declaratives which can be interpreted as rejections or as rejecting questions (RQs, roughly 'disbelieving' questions) shows that RQs differ from rejections in maximum f0 and f0 excursion of the accented syllables, and in the duration and peak alignment of the pre-final syllable. The differences match those that have been observed in comparisons of 'ordinary' declarative questions and assertions. The study included a manipulation of semantic focus (verb/object).

1 Introduction

In Swedish main clause declaratives, the negative marker *inte* can optionally front to the sentence-initial, pre-verbal position in certain contexts, see (1a), which shows the normal position of the negation, and (1b), which shows the same sentence with fronted negation.

(1) a. Anna målar inte Maja   b. Inte målar Anna Maja
      Anna paints not Maja   not paints Anna Maja

According to Petersson (2008), (1b) can be read either as a rejection of a previously asserted claim that Anna is painting Maja, or as a type of question with the rough English translation *Surely Anna is not painting Maja?*[1] Following Seeliger (2015), we call the question reading of (1b) *rejecting question* (RQ). RQs are felicitous in a situation where the speaker had assumed the negative proposition denoted by the declarative ($\neg p$) to be true, but where there is contextual evidence for the corresponding positive proposition (p). For instance, (1b) might be uttered in a situation where the speaker thought that Anna would be painting Michael rather than Maja but now sees Maja entering Anna's studio. In a way, the speaker wishes to 'reject' what s/he sees. Still s/he expects an answer from the addressee about what the true state-of-affairs is, hence the term *rejecting questions* (see Seeliger & Repp, 2016 on the pragmatics and semantics of RQs).[2] (1b) used as a RQ differs from its use as a rejection in that it does not outright reject a proposition p (i.e. *Anna is painting Maja*), but only indicates that the speaker had assumed $\neg p$ to be true and would prefer to keep this assumption.[3]

---

[1] There is at least one other prominent reading of sentences with fronted negation, namely that called *additive negation* by Lindström (2007). This reading is outside the scope of this paper.

[2] Seeliger and Repp (2016) also provide experimental evidence from an acceptability judgement study which confirms the felicity of sentences like (1b) with an RQ reading in the context described in this paragraph.

[3] Sentences like (1a) also have (at least) two readings: as negative assertions and as so-called *negative declarative questions* (NDQs). NDQs are felicitous in a situation where there is contextual evidence for $\neg p$ but

In view of the fact that sentences like (1b) are principally ambiguous between a reading as rejection and a reading as RQ, the issue arises of how the two speech act types are disambiguated. One option is to insert a modal particle: if the particle *ju* is placed after the verb, (1b) is unambiguously marked as a rejection; if the particle *väl* is placed after the verb, (1b) is unambiguously marked as a RQ. However, the use of modal particles is optional. This paper aims to answer the question if, and if so how, the two readings of a sentence like (1b) are disambiguated by the means of prosody.

2 Question intonation in Swedish

Despite the rejective meaning component that RQs have, it seems reasonable to assume that RQs pattern prosodically with at least some types of questions whereas rejections do not. Previous studies on the intonation of Swedish questions have not always been very precise about reporting what type of question (*yes/no-*, *wh-*, declarative question) they investigated. The following overview is restricted to studies on questions without an interrogative syntax since RQs have a non-interrogative syntax.

Gårding (1979) investigated sentence fragments like *några långa nunnor* ('some tall nuns') with statement vs. question intonation. If uttered as questions, these fragments were characterised by a raising of the local pitch peaks (= raising of 'top line'). The bottom line was largely the same as in statements. There were bigger pitch excursions on lexical accents in questions, and questions were characterised by a suppression of downstep, i.e. a slower utterance-global pitch decrease. An utterance-final rise in pitch – which in languages like English or German often is present in declarative questions (von Essen, 1966; references in Ambrazaitis, Niebuhr & Buanzur, 2015) – was not observed.

House (2003) found in a perception study that declarative sentences (e.g. *Han kan tänka sig åka bil* ('He can imagine himself driving a car')) were more likely to be perceived as questions if the utterance-final lexical accent (i.e. on *bil*) peaked later and higher. While this is no evidence for a final rise in the narrow sense, one could speculate that the latest and highest peak in the experiments was perceived as a simple rise (i.e. the fall was too short to be heard). Another contributor to question interpretation was the length of the pre-final syllable (i.e. *ka*) – the longer this syllable was the more likely listeners were to interpret the declarative as a question. House assumes that listeners interpret this lengthening as a hesitation, and that a hesitation is indicative of uncertainty and non-assertiveness, both of which are more compatible with questions than with assertions.

Strömbergsson, Edlund, and House (2012) extracted 600 Swedish questions (*wh-*, *yes/no-*, alternative, and 'other' questions) from a conversational corpus and divided the average f0 in the first half of each question by the average f0 in the second half, as an approximation of whether the questions had falling or rising intonation. They found that *yes/no-*questions generally were falling, while *wh-*questions generally were rising. It is not quite clear whether RQs would have been categorized as *yes/no-*questions or as 'other' in Strömbergsson's et al. classification, but since they are definitely not *wh-*questions, we do not expect RQs to be rising.

---

the speaker had no assumptions or assumed *p* to be true. For instance, the NDQ in (1a) might be uttered in a situation where the speaker thought that Anna would be painting Maja but now sees Michael entering Anna's studio. Thus, both (1a&b) are so-called *biased questions*. They express that the context is of a certain kind (evidential bias) and that the speaker had certain previous assumptions (epistemic bias; cf. Sudo, 2013; Seeliger & Repp, 2016).

With respect to specific prosodic characteristics of rejections, there is no dedicated previous literature. However, Myrberg (2013) compared the prosody of narrow and broad focus in assertions vs. corrections. The latter are a subtype of rejections. Myrberg found prosodic differences between narrow and broad focus but she did not find differences in focus marking in assertions vs. corrections. For instance when *Havren har kokat färdigt.* ('The oat is done boiling.') was uttered as a response to a question eliciting narrow focus in an assertion ('What is done boiling?') it was realized with the same prosody as when it was uttered as a response to a question eliciting narrow focus in a correction ('Are the eggs done boiling?).

In sum, if we are correct in claiming that RQs ultimately are questions, we may expect their prosodic characteristics to pattern with those of the declarative questions studied by Gårding (1979) and House (2003) and with the *yes-no* questions studied by Strömbergsson et al. (2012). We furthermore hypothesise that rejections are similar enough to assertions in terms of prosody that the experiments we just described and our experiment essentially have an identical baseline of comparison.

3 Experiment

We conducted a production experiment in which naïve speakers produced rejections and RQs in quasi-natural conversations, where speakers judged themselves what kind of speech act they were carrying out, on the basis of the context for each target utterance and on the basis of the punctuation following the target utterance (full stop vs. question mark).

3.1 Method
3.1.1 Participants
9 native speakers of Swedish took part in the experiment (mean age 25, range 20-34). All but one of the participants were from Södermanland or Uppland (i.e. speakers of East Swedish in the classification of Bruce & Gårding, 1978), one speaker was from West Sweden. All speakers were female. The restriction to female speakers was motivated by the exclusion of sex as an additional factor. The sex of the speaker has been found to be relevant in emotional speech (e.g. Bachorowski & Owen, 1995; Scherer, Banse, Wallbott & Goldbeck, 1991; Repp, 2015). RQs and rejections plausibly are emotional, but we were not interested in the special role of the speaker's sex with respect to these speech acts. The participants gave informed consent and received monetary reimbursement.

3.1.2 Design and Materials
The experiment employed a 2×2 design with the factors SPEECHACT (rejection/RQ) and FOCUS (object/verb). The materials consisted of 8 experimental items, each in the four experimental conditions, 16 fillers with two conditions each (*wh*-questions and exclamatives), and three practice items. Each item started with a scene-setting passage and was followed by a dialogue consisting of two turns, see (2). The target utterance – a declarative with fronted negation – was the first utterance of the second speaker, speaker B, in (2) lit. *not paints Anna Maja*. It was a reaction to a claim of the first speaker, in (2) *Anna will paint Maja soon*. The claim provided evidence for a positive proposition, which B rejected or posed a RQ for with her target utterance. The post-target utterances of B clarified whether the response was intended as a rejection or a RQ. If the target utterance was intended as a rejection, B gave reasons for why A's claim was not true (e.g. *You should know that Anna only ever paints men)* and offered a correction *(She is painting Mikael))*. If the target

utterance was intended as a RQ, B gave reasons for why A's claim was unlikely to be true (e.g. *Anna promised to paint Mikael. So she is probably painting him.*). Rejections and RQs were marked orthographically by a full stop and question mark, respectively.

In addition to the speech act differentiation, the post-target utterances of speaker B also served to determine the location of narrow focus in the response. Focus was introduced as an experimental factor for two reasons. First, since rejections and RQs involve a polarity contrast with the previous utterance, we might expect that the finite verb always receives a focus accent (e.g. Höhle, 1982) and other focus marking is suppressed. We wished to explore this possibility. Second, the two focus conditions differ from each other in the (expected) distance of the main accent from the final boundary tone, which might be instructive with respect to the presence of a final rise. (2) is an example for object focus. The third utterance of speaker B introduces an explicit alternative to the direct object (e.g. *Hon målar Mikael* in the rejection condition in (2)). The verb focus conditions contained an alternative for the verb (e.g. *teckna* 'to draw').

(2) Sample item, Object focus

Context description: Ett samtal om Anna som är konstnär. Hon målar ett porträtt just nu.
'A dialogue about Anna, who is an artist. She is painting a portrait at the moment.'

Speaker A Anna ska ju måla Anna snart. Jag är spänd på att se porträttet.
Anna will MP paint Anna soon I am excited on to see portrait.DEF
'Anna is going to paint Maja soon. I'm looking forward to seeing the portrait.'

Speaker B Inte målar Anna [Maja]F ?/.   = ***target***
not paints Anna Maja

    Rejection: Anna målar ju alltid bara män. Hon målar Mikael.
'Anna isn't painting Maja. She only ever paints men, as you should know. She is painting Mikael.'

    RQ: Hon lovade ju att måla Mikael. Hon borde väl måla honom?
'Surely Anna isn't painting Maja? As we both know, she promised to paint Mikael. Surely she should be painting him?'

All words used in the target sentences had lexical accent 2 and were disyllabic.

3.1.3 Procedure
All participants saw and read all items in all conditions plus filler and practice items, one by one. There were three different pseudo-randomized orders of presentation of the experimental and filler items (experimental items were not allowed to occur back-to-back), with three participants assigned to each order. The experiment was run using the software *Presentation* (Neurobehavioral systems). Participants were told that they were to vocally enact a dialogue between two speakers, and that they were to take the role of the second speaker. Each item was presented visually on a computer screen and acoustically via headphones. A male voice read the context description aloud, and a female voice read the part of Speaker A. Participants were instructed to thoroughly read the part of Speaker B before they recorded it. They were free to repeat their recording until they were satisfied. Every experimental session was divided into four blocks with short distractor tasks in-

between. An experimental session lasted about 45 minutes. The recordings were made in a laboratory at the Humboldt University of Berlin.

3.2   Results

Two out of 288 recordings had to be discarded because of disfluency and a technical problem. The remaining 286 items were annotated in PRAAT (Boersma & Weeninck, 2015): Each item was split into its syllables (allowing per-syllable measurements). Syllable boundaries were annotated at the onset of short consonants and in the middle of long consonants (of which there were only sonorants).

Figure 1 shows a time-normalized curve (ProsodyPro, Xu, 2013) of the contours in the four conditions. We can see an influence of SPEECHACT on f0: f0 tends to be higher in RQs on the object peaks and on the second peak of the verb and/or first peak of the subject. FOCUS effects can be seen on the pitch peaks of the focussed words: focus correlates with higher f0.
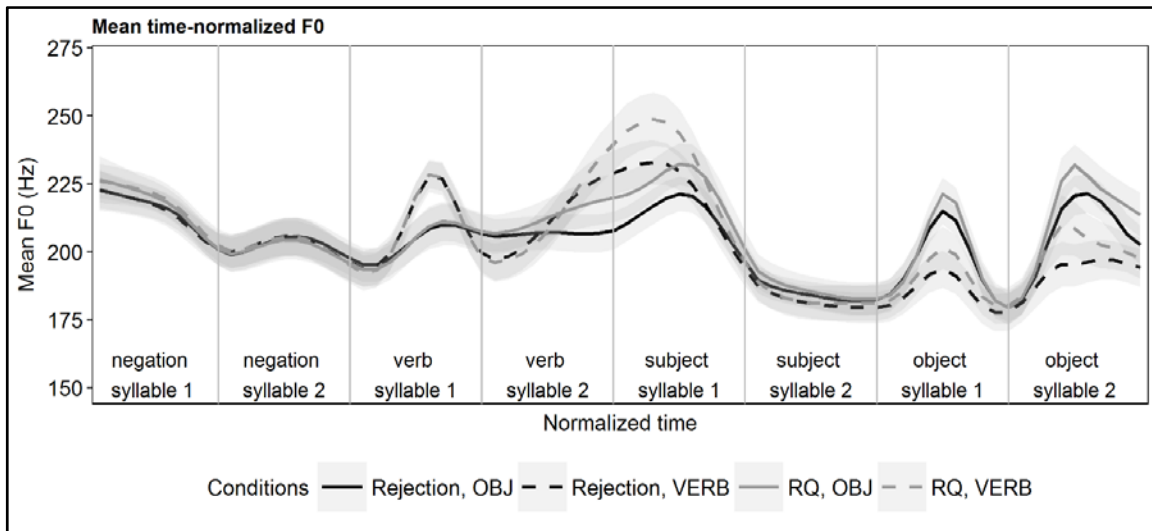

Figure 1. Mean time-normalized f0 (10 measurements per syllable) for the four experimental conditions, averaged over items and participants.

To quantify these results, we fitted linear mixed models (R package `lme4`, Bates, Mächler, Bolker & Walker, 2015) for each syllable for the following acoustic parameters: maximum and minimum f0 ($f0_{max}$, $f0_{min}$) in *Hz*, f0 excursion ($f0_{exc}$) in semitones, duration (*ms*, logarithmized), intensity in dB, and pitch peak position in *ms* from the onset of the syllable. The fixed factors were SPEECHACT (contrast coding: +1 for RQ, -1 for Rejection) and FOCUS (contrast coding: +1 for verb focus, -1 for object focus). Random factors were participants and items. For most measures the best models included random intercepts for participants and items, and random slopes per participant for FOCUS.

The statistical analysis revealed that there were significant effects of SPEECHACT on $f0_{max}$ and $f0_{exc}$ in the verb and object regions (see Table 1), on duration of the pre-final syllable (*b* = 0.02, *se* = 0.009, *t* = 2.5), and a marginal effect on the pitch peak position of the pre-final syllable (*b* = 4, *se* = 2, *t* = 1.99). $f0_{max}$ on the syllables carrying the lexical accents was higher in RQs than in rejections. $f0_{exc}$ was larger in RQs in all syllables except those of *inte* and the second syllable of the verb. The pre-final syllable was longer and the peak position of its

pitch accent was later in RQs.[4] FOCUS had significant effects on $f0_{max}$ in all syllables except those of *inte*, on $f0_{exc}$ of the focussed words and the first syllable of the subject, and on the duration of the focussed words and the first syllable of negation *inte* (Table 1). All these measures were higher on focussed than on non-focussed syllables (except for the duration on *inte*, which was shorter). Pitch peaks were later on the second syllable of the verb in cases of verb focus, on the first syllable of the object in cases of object focus, and earlier on the first syllable of the subject in cases of verb focus. There were no other significant effects of SPEECHACT and FOCUS and no interactions. In addition to the acoustic analyses, all target utterances were annotated for the final boundary tone. One utterance ended in a H% tone. All others ended in L%.

Table 1. Per-syllable model parameters. *b*: effect size; *se*: standard error; *t*: *t*-value. Effect sizes are shown for RQs in the SPEECHACT conditions (reference level = rejections) and for verb focus in the FOCUS conditions (reference level = object focus). We take *t*-values above 2, and below -2 to be significant; "n.s." = not significant.

	word	*negation*		*verb*		*subject*		*object*	
	syllable	1	2	1	2	1	2	1	2
SPEECHACT $f0_{max}$ (Hz)	*b*				5.3	7.2		3.8	6.0
	se				1.4	1.4		1.2	1.7
	t	n.s.	n.s.	n.s.	3.7	5.1	n.s.	3.3	3.6
$f0_{exc}$ (semitones)	*b*			0.2	0.4	0.4		0.3	0.3
	se			0.1	0.1	0.1		0.1	0.1
	t	n.s.	n.s.	2.2	3.2	4.1	n.s.	3.1	2.2
FOCUS $f0_{max}$ (Hz)	*b*			7.9	10.6	8.6	-2.4	-9.9	-11.3
	se			2.3	3.2	3.1	1.1	2.2	3.3
	t	n.s.	n.s	3.5	3.3	2.7	-2.2	-4.5	-3.4
$f0_{exc}$ (semitones)	*b*			0.8	1.1	0.7		-0.8	-0.8
	se			0.2	0.3	0.2		0.2	0.3
	t	n.s.	n.s.	3.5	4.3	3.5	n.s.	-4.5	-2.9
Duration (log)	*b*	-0.04		0.16	0.11			-0.11	-0.08
	se	0.02		0.04	0.02			0.02	0.02
	t	-2.13	n.s.	4.22	4.36	n.s.	n.s.	-4.40	-4.28
Pitch peak position (ms)	*b*				28	-9		-13	
	se				6	3		4	
	t	n.s.	n.s.	n.s.	4.48	-3.37	n.s.	-3.63	n.s.

---

[4] This is difficult to see in Figure 1. Note that Figure 1 is time-normalized, so the length of the syllable is not depicted. If the peak position is calculated in proportion to syllable length, the effect of SPEECHACT disappears and the FOCUS effects are slightly different. This needs closer scrutiny in future research.

3.3 Discussion

The experimental results indicate that there are robust differences between RQs and rejections in Swedish. There were consistent increases in $f0_{max}$ and $f0_{exc}$ of the pitch accents in RQs. These results suggest that in RQs, like in other declarative questions, question marking is distributed across the whole utterance by a rising of the topline (Gårding, 1979). The bottom line did not differ between RQs and rejections, which also matches Gårding's (1979) findings. Furthermore, the pre-final syllable was lengthened, which is compatible with House's (2003) result that a lengthening of that syllable correlates with question interpretation. The tendency for a later pitch peak position of the same syllable was not predicted on the basis of House's (2003) results. In the present experiment, there was a later pitch peak in the pre-final syllable, not in the final one. This difference needs closer scrutiny in future research. It might have to do with the rejective meaning component of RQs (but also see fn. 4). Finally, there was no evidence whatsoever for a final rise. We take this to show that RQs, like Swedish declarative questions (Gårding, 1979) and *yes/no*-questions (cf. Strömbergsson et al., 2012), do not exhibit a high boundary tone / final rise by default.

The results for focus indicate – as would be expected on the basis of previous findings for prosodic effects of semantic focus in Swedish (cf. the overview in Myrberg & Riad 2015) – that accented syllables of focussed words showed a higher $f0_{max}$ and a larger $f0_{exc}$. Duration was longer, and the peak on the second syllable of the verb and on the first syllable of the object were aligned later. These results are not surprising. What is notable, though, is that object focus seemed to receive its 'normal' prosodic marking despite the polarity contrast between the rejection/RQ and the preceding utterance. There were considerable differences between focussed and non-focussed objects even though polarity was in semantic focus as well, that is the speech-act related polarity focus does not seem to interact with other prominence markings in the clause. Such an interaction was observed in other emotional speech acts: the prosodic marking of information structure (focus vs. givenness) in *wh*-exclamatives in German has been found to be reduced in comparison to *wh*-question (Repp 2015). Whether the observed differences in the present experiment are smaller than in sentences without polarity focus must be investigated in future research. A final interesting result is that the speech act difference was only realized on the second pitch peak associated with the verb whereas the focus difference was realized on both peaks. It is an open issue at the moment why the speech act difference did not show earlier and whether this is related to the particular rejective speech acts that we tested.

4   Conclusion

The present study compared the prosody of rejections and RQs. Its results indicate that the markers of question intonation in RQs are the same as those reported in previous studies on other types of Swedish questions with declarative syntax, i.e. the specific prosodic means that mark RQs as questions are shared with other non-lexically marked questions. More production studies are needed which systematically compare RQs to negative declarative questions in different contexts and with different speaker assumptions (see fn. 3), to ascertain if there are specific prosodic effects that are due to the specific speaker assumptions, e.g. there might be an incredulous component in the intonation.

5 References

Ambrazaitis, G., Niebuhr, O., & Buanzur, T. C. (2015). Focal f0 peak shape and sentence mode in Swedish. *Proceedings of the 18th ICPhS, Glasgow, Scotland*. 1-5

Bachorowski, J.-A., & Owren, M.J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science, 6*, 219–224.

Bates, D., Mächler, M., Bolker B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67(1)*, 1–48.

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.04, http://www.praat.org/

Bruce, G., & Gårding, E. (1978). A prosodic typology for Swedish dialects. *Nordic Prosody 1*, 219–228.

von Essen, O. 1966. *Allgemeine und angewandte Phonetik*. Berlin, Germany: Akademie Verlag.

Höhle, T. N. (1992). Über Verum-Fokus im Deutschen. In Jacobs, J. (Ed.), *Informationsstruktur und Grammatik* (pp. 112–141). Opladen, Germany: Westdeutscher Verlag.

House, D. (2003). Hesitation and interrogative Swedish intonation. *Phonum, 9*, 185–188.

Gårding, E. (1979). Sentence intonation in Swedish. *Phonetica, 36(3)*, 207–215.

Lindström, J. (2007). Initial clausal negation - a Nordic areal feature. In Amman, A. (Ed.), *Linguistics Festival* (pp. 31–58). Bochum, Germany: Brockmeyer.

Myrberg, S. (2013). Focus type effects on focal accents and boundary tones. In *Proc. 26th Annual Swedish Phonetics Conference (FONETIK 2013)*. 53–56.

Myrberg, S., & Riad, T. (2015). The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics, 38*, 115–147.

Petersson, D. (2008). Inte, nog och visst i mittfält och fundament. *Nordlund, 29*, 111–153.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Repp, S. (2015). On the acoustics of wh-exclamatives and wh-interrogatives: Effects of information structure and sex of speaker. *Proceedings of the 18th ICPhS*. Paper nb. 319. pp. 1–5.

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion, 15,* 123–148.

Seeliger, H. (2015). "Surely that's not a negative declarative question?" Polar discourses in Swedish, German and English. *Proceedings of Sinn und Bedeutung, 19*, 571–589

Seeliger, H., & Repp, S. (2016). Biased declarative questions in Swedish and German: The syntax of negation meets modal particles (*väl*, *doch* and *wohl*). Submitted.

Strömbergsson, S., Edlund, J., & House, D. (2012). Prosodic measurements and question types in the Spontal corpus of Swedish dialogues. *Proceedings of INTERSPEECH 2012*, 839–842

Xu, Y. (2013). ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis. *Proceedings of TRASP 2013*, Aix-en-Provence, France. 7-10.