Some Questions and Answers about Polish Questions

Jakub Suchojad (jhs212@ruccs.rutgers.edu)
Dept. of Psychology, Rutgers University – New Brunswick
Piscataway, NJ 08854

C. Jane Lutken (c.jane.lutken@rutgers.edu)
Center for Cognitive Science, Rutgers University – New Brunswick
Piscataway, NJ 08854

Karin Stromswold (karin@ruccs.rutgers.edu)
Dept. of Psychology and Center for Cognitive Science, Rutgers University – New Brunswick
Piscataway, NJ 08854

Abstract

Languages differ in how they form questions that are equivalent to English questions such as *who does John think Maria loves*? in that the correct answer is who John *thinks* Maria loves, and not who Maria *actually* loves. Linguists disagree about how Polish makes such inquiries, and to date, no research has investigated how native Polish-speaking adults judge, process or produce these inquiries. In this paper, we investigated the nature of Polish questions via a corpus study, a grammaticality judgment study, and a spoken production study. Taken together, the results of these studies suggest that Polish has several syntactically distinct options for making these sorts of inquiries. Although, at first blush, this seems inconsistent with linguistic theories that argue against syntactic optionality, closer examination reveals that discourse context strongly affects which option is preferred. These findings highlight the importance of considering context, and the pitfall of studying sentences presented in isolation when evaluating linguistic or psycholinguistic claims.

Keywords: language processing; sentence production; syntax; syntax-semantic interface; prosody; discourse; *wh*-questions

Introduction

Contentful questions (henceforth *wh*-questions) that ask about the identity of an agent (who), object (what), location (where), time (when), manner (how) or reason (why) have been the subject of much research by linguists, psycholinguists, and computational linguists because in many languages they appear to involve discontinuous or displaced constituents. In addition, the structure of *wh*-questions varies considerably across languages. Consider the two-clause declarative sentence in (1a) and the two-clause *wh*-question in (1b). Languages like English are said to have long distance *wh*-extraction because, to inquire about the identity of the person that John thinks Maria loves, *who* appears in sentence-initial position rather than at the end of the subordinate clause where *Bill* appears. In the English multi-clause question (MCQ) in example (1b), the initial *who* refers to the person doing the loving, so the entire question is within the scope of the initial *wh*-word.

(1a) John thinks Maria loves Bill**.**
(1b) Who does John think Maria loves [___]?

To ask the equivalent of an English MCQ like (1b), languages like German have a contentful *wh*-phrase medially, the scope of which is marked by an initial, contentless *wh*-phrase as in (2).

(2) German MCQ (example from Dayal, 1994)
Was glaubst du, mit wem Maria gesprochen hat?
What think du with whom Maria spoken has?
'With whom do you think Maria has spoken?'

Although there is considerable disagreement among linguists about the structure of Polish questions, linguists agree that Polish does not allow English-like long-distance extraction from an embedded clause (Lubańska, 2004; Śmiecińska, 2009). Like German, it has been argued that Polish has MCQs, with an initial content-less *wh*-scope marker (WSM), and a contentful *wh*-phrase between the matrix and subordinate clause (Stepanov, 2000; Lubańska, 2004; Śmiecińska, 2009). However, there is disagreement about the exact structure of MCQs. Some linguists argue that the scope marker is *jak* (Stepanov, 2000; Lubańska, 2004) and others argue it is *co* (Śmiecińska, 2009). Further, some linguists (Śmiecińska, 2009) argue that Polish MCQs with true *wh*-scope marking have the structure shown in (3a) with a question-initial scope marker and the complementizer *że* ('that') preceding the embedded clause, whereas others (Stepanov, 2000; Lubańska, 2004) argue that they can have the structure shown in (3b) with a question-initial scope marker and no complementizer.

(3) Polish MCQs ('Who do you think Mary loves?')
(a) *Co/Jak myślisz, że kogo kocha Marysia?*
WSM think that whom love Mary
(b) *Co/Jak myślisz, kogo kocha Marysia?*
WSM think whom love Mary

In addition to MCQs Polish also appears to be able to express the same semantic content of questions like (1b) by using two sequential questions (SQs), the first one of which begins with *co* (4a) or *jak* (4b).

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

(4) Polish SQs ('Who do you think Mary loves?')

(a)	*Co*	*myślisz?*	*Kogo*	*kocha*	*Marysia?*
	How	think?	Whom	love	Mary
(b)	*Jak*	*myślisz?*	*Kogo*	*kocha*	*Marysia?*
	What	think?	Whom	love	Mary?

Note that despite conveying the same information and being lexically identical, Polish MCQs that lack a complementizer such as (3a) and SQs such as (4a) and (4b) are fundamentally different at the *syntactic* level because SQs involve semantic relationships between sentences, and not syntactic relationships within a sentence.

In addition to the ongoing linguistic controversies described above, to date, no psycholinguistic research has investigated how Polish adults process or produce questions that convey the semantics of (1b) and (2). We conducted a corpus study and two experiments to address the linguistic and psycholinguistic nature of complex Polish questions.

Corpus Study

To determine the frequency of different types of questions in Polish, we conducted targeted analyses of the National Corpus of Polish (Przepiórkowski, 2012). This corpus is a large repository of searchable prose, newspaper articles, journal articles, conversation transcripts, internet texts and other sources aimed at providing a well-balanced representation of Polish. Using the PELCRA (Polish and English Language Corpora for Research and Applications) search engine (Pęzik, 2012), we searched the "balanced" version of the corpus (which contains over 250 million words) for the following words and phrases: *co, jak, co myślisz* ("what think"), *jak myślisz* ("how think"), *co myślisz że* ("what think that"), and *jak myślisz że* ("how think that"). For phrases, the words had to be in the exact order specified, but punctuation, capitalization and additional words at the phrase boundary were allowed. In some cases, words in phrases were split between sentences or questions.

These analyses revealed that, overall, *jak* occurred more often than *co* (776,402 vs. 616,402 times, respectively). In addition, the combination of *jak* and *myślisz* appeared more than twice as often as *co* and *myślisz* (750 vs. 301 times, respectively). When the complementizer *że* was added to the search terms, the triad of *co, myślisz* and *że* occurred six times as often as the triad *jak, myślisz* and *że* (46 vs. 7, respectively). However, this finding must be interpreted with caution because the number of such questions was small.

In summary, although our corpus study sheds some light on linguistic controversies surrounding the structure of Polish questions, several issues remain unresolved. Does our finding that twice as many questions contain *jak* as *co* mean that *jak* is *the* scope marker in Polish, or does the fact that *co* questions are relatively common indicate that both *jak* and *co* are scope markers in Polish? Or should we take seriously that, when multiclause questions do have a complementizer, the scope marker is almost always *co*? Does the paucity of clear MCQs with complementizers indicate that Polish MCQs don't have complementizers? To address these questions, we conducted a judgment study in which participants rated different types of questions.

Experiment 1: Judgment Task

Methods

Participants. Fourteen Polish-speaking adults participated (9 female, 5 male). For all participants, Polish was the only language spoken in the home prior to the participants entering school. All but two participants lived in Poland at the time of testing, with most living in southern Poland. The two participants who lived outside of Poland at the time of testing lived in Poland until they were 21 years old.

Stimuli. Participants silently read 32 short stories (mean length = 83.8 words, range 45-145 words) that were designed to elicit complex questions. After each story, participants rated how well 6 written options "fit" the story. The 6 options were SQs with *jak* (6a) or *co* (6b), MCQs that lacked a complementizer (MCQ -comp) and had either the WSM *jak* (6c) or *co* (6d), and MCQs that had the complementizer *że* (MCQ +comp) and either *jak* (6e) or *co* (6f).

(6) Example of six question types (All translate as 'Kasia, what do you think we should see?')

a. SQ *jak*: *Kasia, jak myślisz? Co powinnyśmy zobaczyć?*
b. SQ *co*: *Kasia, co myślisz? Co powinnyśmy zobaczyć?*
c. MCQ -comp *jak*: *Kasia, jak myślisz, co powinnyśmy zobaczyć?*
d. MCQ -comp *co*: *Kasia, co myślisz, co powinnyśmy zobaczyć?*
e. MCQ +comp *jak*: *Kasia, jak myślisz, że co powinnyśmy zobaczyć?*
f. MCQ +comp *co*: *Kasia, co myślisz, że co powinnyśmy zobaczyć?*

The second (contentful) *wh*-word varied, with *co* ('what'), appearing after 11 stories, *gdzie* ('where') appearing after 7 stories, *kiedy* ('when') appearing after 5 stories, *kto* ('who') appearing after 4 stories, *kogo* ('whom') appearing after two stories, and *ile* ('how much/many'), *która* ('which one') and *jak* ('how') questions each appearing after one story.

Roberts (2012) has argued that every exchange is made up of various "Questions Under Discussion" which can be either closed (if the answer is known) or open (if the answer is unknown). Lutken & Legendre (2022) suggested that SQs are better when the person asking the question does not know the answer to the second/embedded question ("open" contexts), and MCQs are better when the person asking the question does know the answer to the second/embedded question ("closed" contexts). Thus, it is possible that both SQs and MCQs are acceptable in Polish, but in different contexts.

For this reason, we designed the stories such that the answer to the second/embedded question was unknown to the speaker in 20 stories ("open" context stories) and known to the speaker ("closed" context stories) in 12 stories. For

example, (7a) is an open context story because the person asking the question (Gabrysia) does not know who Kalina will choose to invite to the musical. In contrast, (7b) is a closed context story because Jacek knows where the amusement park ride is.

(7) Examples of open and closed context stories

a. Open context story

Gabrysia and Jarek are going to a musical at the Grand Theatre and have two extra tickets. One of them is for their friend Kalina, but they don't know whom else to invite. They decide to let Kalina choose who they should invite. Gabrysia calls Kalina, explains the situation, and asks Kalina ___

b. Closed context story

Jacek takes his three sons to a water park. They try to find their way to their favorite attraction. The two youngest boys start arguing at a crossroads: should they go right or left? Stefek thinks he remembers that the way to the ride is to the left, while Andrzej thinks they should go right. Jacek notices a sign saying they should go right. The youngest, Krzyś, is just starting to read. Jacek leans down, points to the sign, and asks Krzyś ___

Procedure. The web-based experimental testing platform FindingFive (Finding Five Team, 2019) presented the stimuli and recorded participants' responses. Stories were presented in a written format on the screen, followed by a question to be rated (in bold). Under the question there was slider-style widget, the extreme ends of which were labeled 'does not fit' (1) and 'fits' (5). After a participant rated the fit of a question, that question disappeared, and another version of the question appeared. The story stayed on the screen, until all six versions of the question were presented. The six versions of questions (6a-f) appeared in random order across the 32 stories. The task was self-paced, and halfway through the experiment, the participants took a break and answered questions about themselves and their linguistic background. The Rutgers University Human Subject Institutional Review Board approved the study, and participants received $15 for participating.

Results

Due to a coding error, an open context "where" story was excluded. Participants' ratings for the remaining 186 questions were normalized by participants, and these normalized ratings were analyzed in JASP v. 0.18.3 (JASP Team, 2024) using Bayesian ANOVAs with participant as a random variable.

To assess the impact of syntactic structure, a 3 (SQ/MCQ -comp/MCQ +comp) x 2 (Jak/Co) ANOVA was performed. This revealed that, in the best-fitting model ($BF_{10} > 1x10^{+48}$), there were main effects of *jak/co* (with *jak* questions being rated higher, $BF_{inc} = 2.5x10^{+14}$) and question type ($BF_{inc} = 2.5x10^{+14}$), and the interaction between these factors ($BF_{inc} = 398$). Post hoc analyses revealed that the main effect of question type was due to SQs being rated highest (mean Z-score = .284), then MCQ -comps (mean Z-score = .018), with MCQ +comps (mean Z-score = -.301) rated lowest (BF_{10} for all pair-wise comparisons > $2x10^{+5}$). As depicted in Figure 1, the interaction between question type and choice of scope marker was due to SQs and MCQ -comps with *jak* being rated higher than those with *co* (both $BF_{10} > 3x10^{+7}$), whereas the ratings were the same for MCQ +comps that had *jak* and those that had *co* ($BF_{10} < 1$).

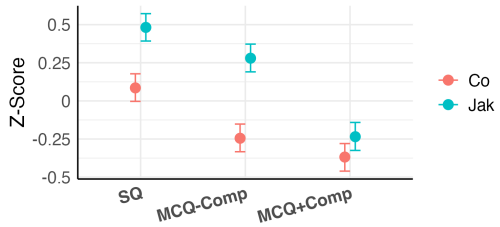

Figure 1. Normalized ratings for questions with *co* and *jak* by question type (error bars = 95% credible intervals)

Individual subject analyses revealed that participants differed in their relative preferences for the three types of questions (see Figure 2). Seven participants rated SQ and MCQ -comp questions to be equally good, with all 7 rating both types of questions to be better than MCQ +comp questions. Three participants rated SQs to be substantially better than both types of MCQs, with all 3 participants rating MCQ +comp and MCQ -comp questions to be equally bad. One participant rated both types of MCQs to be equally good and better than SQs. Finally, three participants rated all three types of questions to be equally good.

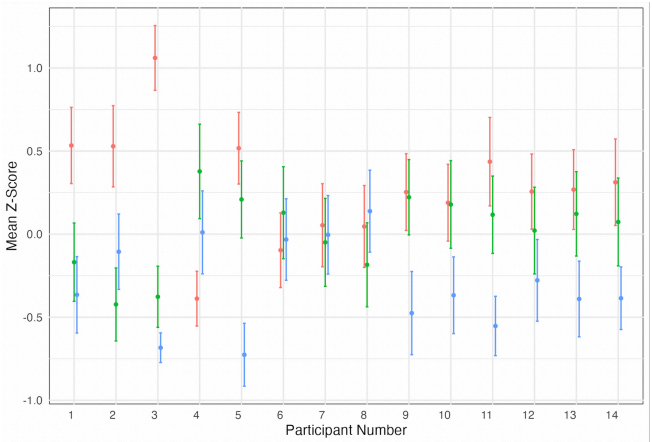

Figure 2. Individual participants' rating for questions that are SQ (red), MCQ -comp (green) and MCQ +comp (blue). (error bars = 95% credible intervals)

To assess the role of discourse context, a 3 (question type) x 2 (open/closed context story) ANOVA was performed. In the best fitting model ($BF_{10} = 1.76 \times 10^{+35}$), there were main effects of context (with questions being rated higher after open context stories, $BF_{inc} = 2.99 \times 10^{+4}$) and question type as

described above (BF_{inc} > $1x10^{+50}$), and the interaction between these factors (BF_{inc} = 5.15 $x10^{+4}$, see Figure 3). Post hoc analyses revealed that this interaction was due to SQs receiving decisively higher ratings after open than closed context stories (BF_{10} = 3.00 $x10^{+3}$), MCQ -comps receiving higher ratings after open context stories (BF_{10} = 12.4), and MCQ +comps received somewhat worse ratings after open context stories (MCQ +comp BF_{10} = 5.90).

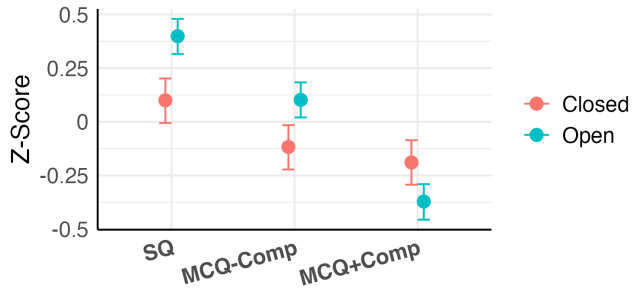

Figure 3. Ratings for questions appearing after open and closed context stories (error bars = 95% credible intervals)

In summary, overall, our participants strongly preferred SQs to MCQ -comps which in turn they preferred to MCQ +comps. However, the situation may be more complicated. Individual subject analyses revealed that only three of the 14 participants robustly preferred SQ questions to MCQs, whereas 7 participants rated SQ and MCQ -comp to be equally good. Indeed, only one (participant 3) robustly exhibited the SQ > MCQ -comp > MCQ +comp preference, with all other participants rating two or three of the constructions as being equally good.

Overall, our participants gave higher ratings to questions that appeared after open context stories than closed context stories. However, the effect of context varied for the three types of questions, with SQs and MCQ -comps getting higher ratings after open context stories and MCQ +comps getting higher ratings after closed context stories. Thus, the apparent coexistence of multiple different ways of making the same inquiry appears to partially reflect the effect of discourse context, and not true optionality.

Our judgment study has several limitations. First, participants' ratings of questions may have been affected by the fact that the questions were written. Specifically, participants' higher ratings for SQs could reflect SQs being more orthographically salient because they contain two question marks and both *wh*-words are capitalized. Similarly, participants' strong dislike of MCQs with complementizers could reflect that they are longer than the other two types of questions, and therefore they are harder to read.

A second concern has to do with the fact that, after each story, the participants rated all 6 possible forms of the question. This may have biased the participants to rate, not simply rate the absolute goodness of a question, but to compare the relative goodness of that question with those that preceded it. It is also possible that repeatedly rating the same types of questions resulted in participants' "losing" their ability to rate them, a phenomenon referred to as "syntactic satiation" (Stromswold, 1986; Snyder, 2000; Snyder, 2022).

Lastly, it is possible that, if participants had been able to, they would have asked questions with a different structure than the 6 options we gave them. To address these concerns, we conducted a second experiment in which participants read stories and then were free to verbally respond with a question that had whatever form they wished.

Experiment 2: Free Response Study

Methods

Participants. Twelve adult Polish speakers participated. Of these, 10 also participated in Experiment 1. For those who participated in both experiments, the average interval between the two experiments was 30 days, with no interval being shorter than 21 days.

Procedure. The stories were the same as those used in Experiment 1, but rather than rating different questions, participants were told to 'ask whatever question they thought fit best.' Stories were presented on the screen in the same fashion as in Experiment 1. However, at the end of a written story, there was a fill-in-the-blank space followed by a question mark. A small microphone icon appeared beneath the story. Participants clicked on the icon to begin and to end recording. Participants were allowed to re-record their answers until they were satisfied with their response.

Results

Textual analyses of free responses. The audio files were transcribed by a native Polish speaker. If a participant recorded more than one question for a story, we analyzed their final response. As depicted in Figure 4, of the 384 productions, 117 were potentially SQ or MCQ because they

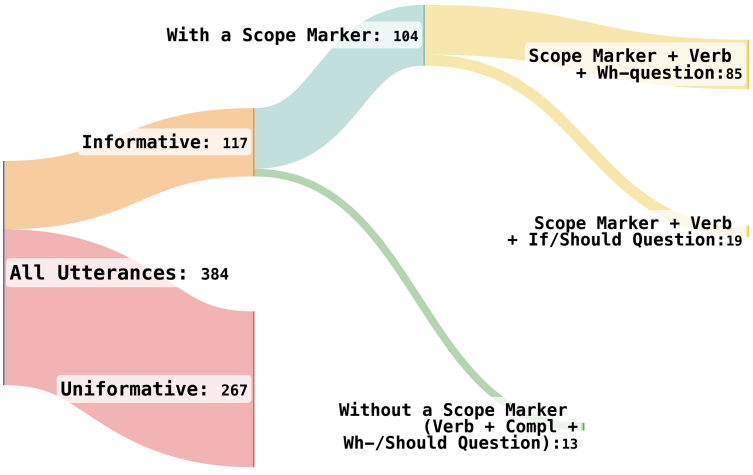

Figure 4. Distribution of Free Responses

had at least two sentential clauses (i.e., two verbs), the clause

of the first verb was a "thinking" verb, and the second clause contained a *wh*-word, a complementizer, a modal verb conveying possibility (similar to the English modal verb *could*) or the Polish particle *czy* (which can introduce interrogative questions in a fashion similar to a *wh*-word). All other utterances were monoclausal, sentence fragments, or otherwise considered uninformative.

Eleven of the 12 participants asked informative questions. Of the 117 informative questions, 101 contained *jak*, three contained *co*, and 13 had no discernable scope marker. Only 14 of the 117 questions had a complementizer, and in all cases the complementizer was *że*. Of the 14 questions with a complementizer, 11 had no discernable scope marker and 3 had the scope marker *jak*.

Acoustical analyses of responses. Because SQs and MCQs that lack complementizers don't differ lexically (compare, for example, 6a and 6c, and 6b and 6d), we examined whether they differed prosodically. The audio files for informative utterances were analyzed using Praat version 6.4.04 (Boersma & Weenink, 2024). The files were trimmed so that they began with first transcribed element, including discourse markers (e.g., 'well' or 'so'), discourse particles (e.g., 'um' or 'like'), the names of characters in the stories being addressed (e.g., 'Kasia'), and repeated or stuttered words. F0 was normalized by participant, and the prosodic contours shown in Figures 5 and 6 depict the mean normalized F0 with error ribbons corresponding to 1 standard error. The F0 data were quite sparse by 3500 ms, and so the prosodic curves in these figures are truncated at 4000 ms.

A native Polish speaker and four Polish-naïve linguists listened to each informative production and judged whether it was prosodically a single sentence (i.e., a potential MCQ) or two sentences (i.e., a potential SQ). The Fleiss's kappa for the 5 raters was .38 ($p < .00001$). The classification of questions as SQ or MCQ was based on majority decision.

Forty-nine questions were deemed SQ and 68 were deemed MCQ. As depicted in Figure 5, the F0 curves for questions

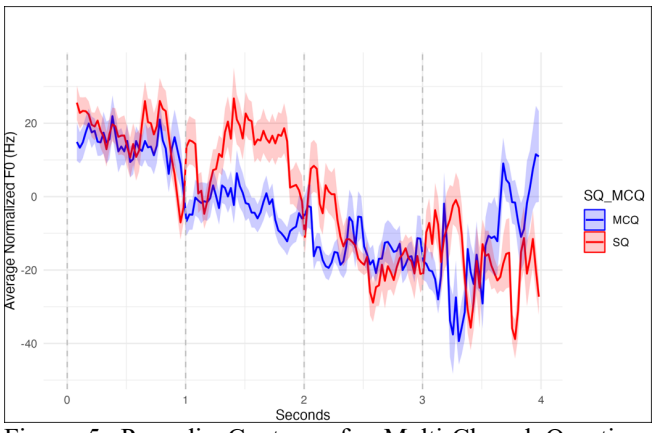

Figure 5. Prosodic Contours for Multi-Clausal Questions (blue) and Sequential Questions (red).

classified as SQs (red line) and MCQs (blue line) initially overlapped considerably with both precipitously dropping 30-40 Hz at about 1000 ms. At 1000 ms, the SQ and MCQ F0 curves clearly diverged. For SQs, the F0 rebounded 30 Hz, and remained elevated until 1750 ms. This overall pattern is consistent with productions that were classified as SQs being two distinct questions, with the second question beginning around 1250 msec. In contrast, for questions classified as MCQs, the F0 dropped a bit later (at about 1000 ms) and continued to steadily decline with no evidence of a subsequent F0 peak, consistent with productions classified as MCQs being a single multi-clause sentence.

As depicted in Figure 6, during the first 2000 ms, the F0 prosodic contours of the 64 informative questions asked after open context stories (red line) differed markedly from the 53 informative questions asked after closed context stories (blue line). Questions asked after open context stories (red line) had consistently higher mean F0s during the first 750 ms. Between 1000 and 1250 ms, for open context questions (red line), there was a sharp, narrow F0 trough followed by a pronounced F0 peak at 1500ms. In contrast, for questions asked after closed context stories (blue line), there was no sign of an F0 peak between 1000 and 2000 ms. Indeed, there was a hint of a dip in F0 at around 1500 ms.

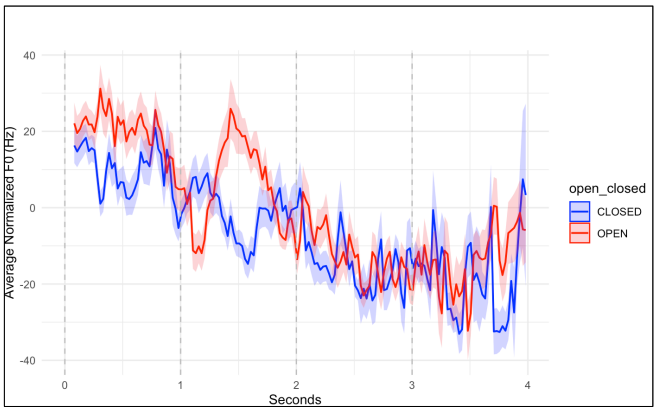

Figure 6. Prosodic contours for questions said after open context stories (red) and closed context stories (blue)

If we compare the prosodic contours for the questions classified as SQs (red line, Figure 5) and questions asked after open contexts (red line, Figure 6), in both cases, there was a clear F0 dip at about 1000 ms, followed by an F0 peak at about 1500. This prosodic similarity is consistent with our participants asking (mostly) SQs after open context stories. If we compare the prosodic contours for questions classified as MCQs (blue line, Figure 5) with questions asked after closed contexts (blue line, Figure 6), in both, cases the F0 declines fairly steadily over the course of the utterance. This prosodic similarity is consistent with our participants asking (mostly) MCQs after closed context stories.

Discussion

The complexity of complex Polish questions. We began this paper by asking how Polish forms questions that are equivalent to English questions such as *who does John think Maria loves*? where the correct answer is who John *thinks*

Maria loves, and not who Maria *actually* loves. Although linguists generally agree that Polish has MCQs with a question-initial scope marker and a medial *wh*-phrase, they disagree about the structure of MCQs. Some argue the scope marker is *jak* (Stepanov, 2000; Lubańska, 2004) and others that it is *co* (Śmiecińska, 2009). Similarly, some argue that only MCQs with complementizers have *wh*-scope marking as seen in language like German (Śmiecińska, 2009) and others argue that complementizers are not required (Stepanov, 2000; Lubańska, 2004). Our findings that, *jak* questions were more common in a corpus study, were rated better in a judgment task, and were produced more often in a free response task suggest that the Polish scope marker is *jak.* Similarly, our findings that MCQs with complementizers were very rare, received very low ratings, and were rarely produced suggest that Polish MCQs don't require complementizers.

Recall that, in Polish, the semantic content of an MCQ can be expressed by two SQs. Indeed, overall, participants rated SQs to be much better than MCQ -comps which were in turn rated to be much better than MCQ +comps. Although these ratings could be taken as evidence that Polish doesn't have MCQs of *any* sort, in our production study, participants asked similar numbers of SQs and MCQ -comps (42% and 46%, respectively), and avoided asking MCQ +comps (12%).

Participants' ratings of questions in the judgment study and the prosodic characteristics of the questions they produced in the free response study indicate that Polish has both SQs and MCQs, but they occur at different rates in different settings. When the discourse context is clearly open, people ask SQs, and when the discourse context is clearly closed, they are willing to ask MCQs with complementizers.

Why might this be? We believe it reflects that participants prefer syntactically simple constructions, but this preference can be modulated by the semantics of the discourse context. All else being equal, participants prefer the simplest possible construction. Because SQ are just two simple, monoclausal questions, they are always preferred to MCQs. Furthermore, because SQs are more felicitous after open than closed discourse contexts, the SQ preference is particularly apparent in open contexts. MCQ +comps are the most complex. Thus, they are the least preferred, although they are better in closed contexts where they are the most pragmatically felicitous. Consistent with people avoiding complex structures, in the rare cases when our participants *did* ask MCQ +comps, they omitted the scope marker 80% (11/14) of the time. Perhaps if we had included MCQ +comps without scope markers in our judgment study, MCQ +comps would have received higher ratings. Complexity-wise, MCQ -comps are between simple SQs and MCQ +comps and, consistent with this, they received middling ratings.

In summary, the results of our three studies indicate that Polish has both SQs and MCQs but they don't necessarily have the structure described by linguists. Furthermore, SQs and MCQs are used in different contexts and, thus, some of the disagreement about the structure of Polish questions likely reflects context-dependent alternatives.

Beyond our findings' implications for the linguistic and psycholinguistic nature of Polish questions, our research has broad theoretical and practical implications for scientists who study language.

Task matters. The task used to assess language matters. Although our participants greatly preferred SQs to MCQs in a judgment task, when allowed to pose questions however they liked, they produced them at similar rates, *and* they produced constructions that have not been attested in the linguistic literature. Had we only used a judgment task, we might have concluded that Polish doesn't have MCQs.

Modality matters. Studying spoken language is harder than studying written language. But "real" language is spoken, and written language is an imperfect, incomplete rendition of spoken language. Thus, the modality of a language task can critically affect the results obtained and hence one's conclusions about language. Such was the case with our production task. If our participants had typed their questions rather than saying them aloud, or if we had only analyzed transcriptions of their productions, we might have thought their MCQs were SQs. The same is true for judgment tasks because a sentence may be perfectly acceptable if written, but unacceptable if said with the wrong prosody, and vice versa. Further, because prosody can change the meaning of a sentence or make it easier (or harder) to understand, one may reach different conclusions about the parser from written versus spoken comprehension studies (Cohen et al., 2001). Likewise, although preferable to analyzing written texts, the lack of prosodic information means that even if one is analyzing transcriptions of spoken language, one must proceed with caution.

Individuals matter. We must recognize that people are not all the same. As a group, our participants decisively preferred SQs to MCQ -comps to MCQ +comps, yet half of our participants rated SQs and MCQ -comps to be equally good and, in our production task, some participants asked more SQs than others. Why these differences? Do people have different micro-grammars? Different parsers? Or perhaps it reflects the subjective nature of contexts – one person might interpret a context as open (and hence produce an SQ) and another interpret it as closed (and hence produce an MCQ).

There is a long, storied history of studying written sentences presented in isolation. While this has led to deep insights about language and language processing, doing so may lead one astray regardless of whether one is conducting corpus, judgment, comprehension, or production studies. From linguistic and psycholinguistic perspectives, even if one believes in strong modularity, the modules of language must interface with and affect one another, and the default/citation form of an utterance in isolation may not be the best in a specific context or with a particular intonation. In summary, to truly understand sentences and the way they are processed and produced, we must study how, when and where the sentences are said.

Acknowledgments

This work was partially funded by the David and Dorothy Cooper Scholarship. Thanks go to the members of the Language and Processing Lab and our participants.

References

Boersma, P. & Weenink, D. (2024). Praat: Doing phonetics by computer. Version 6.4.04 http://www.praat.org/.

Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse processing. *Brain and Cognition*, *46*(1), 73–82.

Dayal, V. (1994). Scope marking as indirect *wh*-dependency. *Natural Language Semantics, 2*(2), 137–170.

FindingFive Team (2019). FindingFive: A web platform for creating, running, and managing your studies in one place. FindingFive Corporation (nonprofit), NJ, USA. https://www.findingfive.com

Korotkova, N. (2012). On alleged *wh*-scope marking in Russian. *Proceedings of the 30th West Coast Conference on Formal Linguistics* (pp. 205-215). Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2818.

Lubańska, M., (2004). *Wh*-scope marking in Polish. *Poznań Studies in Contemporary Linguistics 39*,73-88.

Lutken, C.J. & Legendre, G. (2022). *Wh*-scope marking: Cross-linguistic variation at the semantics-syntax interface. Linguistics Society of America 96. Poster.

Pęzik, P. (2012). Wyszukiwarka PELCRA dla danych NKJP. Narodowsky Korpus Jezyka Polskiego [The PELCRA search engine for the National Corpus of Polish]. A. Przepiórkowski, M. Bańko, R. Górski, B. Lewandowska-Tomaszczyk B (editor). 2012. Wydawnictwo PWN. *253–273)*. Do pobrania. https://dspace.uni.lodz.pl:8443/xmlui/handle/11089/1326

Przepiórkowski, A. (2012). *Narodowy korpus języka polskiego*. Naukowe PWN.

Reis, M. (1996). On was-parentheticals and was w-constructions in German. U. Lutz & G. Müller (Eds.), *Papers on wh-scope marking*. Stuttgart/Tübingen: Universität Stuttgart/Universität.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics*, 5(6), 1–69.

Śmiecińska, J. (2009). *Wh*-scope marking strategies in Polish. Formal Description of Slavic Languages conference. Potsdam

Snyder, W. (2000). An experimental investigation of syntactic satiation effects. *Linguistic Inquiry*, *31*(3), 575-582.

Snyder, W. (2022). On the nature of syntactic satiation. *Languages,* 7(1).

Stepanov, A. (2000). *Wh*-scope marking in Slavic. *Studia Linguistica*, *54*(1), 1-40.

Stromswold, K. (1986). Syntactic satiation. Unpublished manuscript. MIT, Cambridge, MA.