Dimensions of (dis)preference in designing polar answers in American English: A latent class analysis

Ryan Ka Yau Lai[a] & Yan Lashchev[b]
University of California, Santa Barbara
kayaulai@ucsb.edu[a], ydl@ucsb.edu[b]

Abstract

How we answer questions is often affected by whether our response conforms with the bias, or *tilt*, encoded in the question. For example, if we have a 'yes' answer to a negatively-tilted question like *You aren't eating, right?*, we may delay, hedge and explain our answer. We examine these phenomena at scale through the Switchboard Corpus: We determine which aspects of answer design tend to appear together and how this relates to question tilt through latent class analysis. We find three groups of design features that, respectively, challenge assumptions of the question-answer sequence, expand on the answer, and delay presentation of the answer. We also find that answers contradicting the question's tilt are much closer in design to tilt-conforming answers than responses without polarity, though they do disfavour answers that have *none* of the three classes of features. Results support a gradient and multi-dimensional conception of conversational preference.[1]

1 Introduction

Questions are often designed to be biased, or *tilted*, towards certain types of responses (Bolinger 1957, Heritage & C Raymond 2021). For example, *This is true, isn't it?* is tilted towards 'yes', and *This isn't true, is it?* towards 'no'. An answer congruous with the question's tilt promotes solidarity; the opposite answer may threaten it. This is part of a wider phenomenon called *preference* in Conversation Analysis (Pomerantz & Heritage 2012, Nishizaka & Hayano 2015, Pillet-Shore 2017), specifically the *preference for agreement*, a type of *action preference*: Some actions (e.g. answering positively a positively-tilted question) are *preferred actions*, while others (e.g. answering negatively a positively-biased question) are *dispreferred actions*.

Previous research finds that people minimise the face threat in dispreferred responses by designing them to be less direct (Sacks 1987 [2010], Pomerantz 1985). They may **delay** the answer using silence, audible breaths, laughter, or words like *well*, *uh*; **qualify** it using phrases like *I think*, or **explain** the answer. Such answers have *dispreferred turn formats*; by contrast, short and straight answers have *preferred turn formats*. In other words, previous research found that action preference and design preference tend to go together: preferred actions tend to be implemented with preferred turn formats, and vice versa.

Traditionally, these observations come from qualitative analyses of small datasets. Recent quantitative studies both confirm these observations and complicate the picture. Stivers et al. (2009) find that responses that do not really answer the question are produced slower than answers, and tilt-non-conforming answers are slower than conforming ones. Roberts et al. (2015) find that positive answers are only slightly (~55 ms.) faster than negative ones. Robinson (2020a) argues against the claim that 'neutral' yes-no questions, e.g. *Do you have cats?* asked by someone who does not know the answer, prefer 'yes'; instead, both 'yes' and 'no' answers are preferred responses, while conditional ('it depends')-type answers are dispreferred. Kendrick & Torreira (2015) found that longer delays are much more strongly associated with dispreferred turn formats than with dispreferred actions. Kendrick & Holler (2017) found that dispreferred responses to polar questions were 123-165 ms slower than preferred ones (depending on the operationalisation).

Previous studies have not extensively investigated differences between the various strategies for creating

[1] Thanks to Karen Nylund-Gibson and Delwin Carter for help with modelling, Simon Todd for extensive comments on the paper, and John W DuBois, and members of UC Santa Barbara's CEILing group and the California Annual Meeting on Psycholinguistics for additional discussion.

Proceedings of the Society for Computation in Linguistics (SCiL) 2025, pages 166-181.
Eugene, Oregon, July 18-20, 2025

dispreferred turn formats, which may serve different functions and have different relationships with action preference. This may be in part due to sample size limitations, as disentangling the many strategies requires more than the 200 or so question-answer pairs analyzed in previous work (Robinson 2020a, Kendrick & Torreira 2015). This study examines these differences using corpus-based computational methods, leveraging rich annotations available for the Switchboard Corpus (Godfrey, Holliman and McDaniel 1992). Focusing on polar (i.e. yes-no) questions and their answers in American English, we aim to answer:

1. Are there regularities as to how different answer design strategies appear together?
2. If so, how are the different groups of strategies related to action preference?

The first question is answered by sorting answers into classes according to different features of turn design, then examining which features are associated with which classes, using a latent class model (Nylund-Gibson & Choi 2018). The second is answered by predicting class membership from action preference, using tilt-conformity as an auxiliary variable (Asparouhov & Muthén 2014).

2 Data and methodology

2.1 Corpus and extraction of question-answer pairs

This study uses the Switchboard Corpus (Godfrey, Holliman and McDaniel 1992), consisting of American English telephone conversations between strangers on researcher-designated topics. We mainly made use of the annotations made available in XML format through the NXT-format Switchboard Corpus (Calhoun et al. 2010) and the Switchboard dialogue act corpus (SwDA) (Jurafsky, Shriberg & Biasca, 1997), as converted into CSVs in Potts (2011).

The corpus is divided into approximately utterance-sized units called *slash units*. SwDA assigns a dialogue act annotation to each slash unit, e.g. `qy` for polar questions, `ny` for 'yes' answers, etc. Tags are often modified by adding letters followed by `^`, e.g. `^r` means something is a repetition. Unless otherwise specified, when mentioning a tag in this paper, all the modified versions are included. Appendix A lists and defines all the SwDA tags relevant to this paper.

Polar questions were extracted by searching for the tags `qy` and `^g`. For each extracted question, the next turn from a different speaker than the one who produced the question was extracted as the answer. Question-answer pairs where there was a gap of 5 seconds or longer between the question and the answer were excluded, as they are likely to be erroneous. See Appendix B for the treatment of rare edge cases like multiple questions and turn increments. After question-answer pairs were extracted, we determined whether the answer implements a preferred action and detected different answer design features.

2.2 Features of answer design

Before extracting the features of responsive turns, each turn was divided into three parts. The first slash unit to convey the polarity of the answer (generally tagged `ny`, `na`, `aa`, `nn`, `ng`, `ar`, `no`, `am`, `arp`, `nd`) is called the *core* of the answer in this paper. The parts preceding it are *pre-core*, and the parts following it *post-core*. Answers without detectable cores are not considered. An example is given in Table 1.

A	1	# Like Garth Brooks. # /	Question	`qy^d`
B	2	Garth Brooks, {F oh } /	Pre-core	`^h`
	3	yes, # /	Core	`ny`
	4	{D you know } he's fine. # /	Post-core	`sv^e`

Table 1: Examples of pre-core, core and post-core slash units.

Features of the responsive turn considered in this study are divided into two groups: Those preceding the core or concerning the core itself, and those following the core. The following paragraphs describe how the features were extracted. Though many features were extracted based on the literature, only those appearing >5% of the time were included in the final dataset. Full details of the extraction process and excluded features are in Appendix B.

Pre-core/Core features. The **OFFSET** between two turns was calculated by taking the timestamps of the last word of the question and the first word of the answer. Non-linguistic vocalisms at edges of turns are not considered part of the turn in this calculation. This resembles Offset 2 of Kendrick & Torreira (2015). A negative number indicates overlap between the two turns; a positive number indicates a gap.

Fillers and discourse markers were tagged in the corpus (Meteer & Taylor 1995). Features related to these words are detected either directly using those tags, or using the forms of words (since there are missing tags):

- **FILLERS**: either words other than *oh* tagged `{F }`or having the form *uh* or *um*

- **DMOther**: discourse markers other than *oh*, tagged `{D }` or with the forms *well* or *you know*.
- **DMOh**: discourse marker *oh*. It is considered separately as it does not serve to delay the answer, but challenges the question's appropriateness and asserts the answerer's epistemic authority (Heritage 1998, 2005).

Other core-delaying features like breath and laughter were excluded as they did not exceed 5%.

Cores were also tagged for whether they are interjection-type – simple, single-word answers that convey polarity and do not grammatically combine with other words – or non-interjection-type ones (**NonInterj**) (called *type-nonconforming answers* in G Raymond (2003)). Cores tagged `nn`, `ny`, are treated as interjection-type answers, plus words like *right*, *yeah*, *sure*, *probably, certainly* when standalone; the rest are non-interjection-type answers. Non-interjection answers are mostly repetitional (Heritage & G Raymond 2012, Enfield 2019), repeating words and grammatical structures in the question (B: *Well, do you do any recycling?* A: *Uh, we do here.*). Some are transformative answers (Stivers & Hayashi 2010) which indirectly imply the answer (A: *You use your, your company's?* B: *My husband's*, which implies a positive answer, but rejects the presupposition that the company is owned by B).

Finally, we looked for words and phrases expressing qualification or epistemic downgrade (**Downgrade**), i.e. lowering the answer's confidence, before or at the core:

- Adverbs like *probably*, *somewhat*, *sometimes*, *personally*, *maybe*, *perhaps*;
- Modal auxiliaries like *could*, *might*, *may*;
- Degree adverbs like *really*, *so*, *very*, *too*, *usually*, with a negator (e.g. *Uh not really*);
- Epistemic/evidential verbs like *think*, *believe*, *guess*, *know*, *say*, *feel*, and common paraphrases, based on Cappelli (2007) and Thompson (2002);
- Slash units tagged `^h` (hedge).

Extraction was aided by part-of-speech tagging and dependency parses from spaCy (Honnibal & Montani 2017) with a three-stage process: adverbs and modal auxiliaries were extracted from the corpus, those related to epistemic downgrade were manually chosen, and then the corpus was reprocessed to detect the chosen forms, reducing the possibility of missing forms that were mistakenly tagged. Note that some downgraders act as interjection-type answers alone (Stivers 2022: 95).

Post-core features. A post-core has the feature **SamePola** if it contains a polarity-conveying dialogue act with the same polarity as the core. It has the feature **CoreExt** if it contains an extension of the core (with the tag `^e`): these are utterances that repeat or qualify the polarity of the answer, but with more complex expressions than the core (e.g. *Yes, I do*.). A post-core has the feature **Expand** if it has a statement (with tag `sv` or `sd`) without the modification `^e` – roughly corresponding to turn expansions (Ford 2001, Lee 2015) in Conversation Analysis. Such expansions can include explanations and elaborations of the core, twists on the core, etc.

Features for fillers, discourse markers, and downgrade were also extracted for the post-core (other than *oh*, which has no known consistent post-core function). An additional feature extracted for post-core but not pre-core is **ConjBut**, consisting of conjunctions *but* and *(al)though*, because they often present information that contrasts with the polarity conveyed by the core, often in order to qualify it.

Feature	Definition	Location	Example
Offset	Time (sec.) between question and answer	PreC/C	B: Do you have **kids**? / A: **[offset = 1.794s] I** have three.
Fillers	Words like *uh* or *um* that fill pauses	Both	**{F Uh, }** we will be.
DMOh	The discourse marker *oh*	PreC/C	{**F Oh, }** I do.
DMOther	Discourse markers other than *oh*	Both	**{D Well, }** {F uh, } I have thought about it.
NonInterj	Repetitional and transformative answers	PreC/C	B: Is Texas one of them? A: **Texas is not one of them**.
Downgrade	Language for epistemic downgrade	Both	**Probably** not.
SamePolA	Polarity-bearing dialogue act with the same polarity as the core	PostC	No, / **no.**
CoreExt	Extension of the core	PostC	No, / **I'm not. / [sd^e]**
Expand	Statements expanding on the core	PostC	Yeah. /**{F Uh, } I understand. [sv]**
ConjBut	Contrastive conjunctions like *but*	PostC	No, / I don't, / **{C but }** I think I know what it is.
SISR	Self-initiated self-repair	PostC	Yeah, / **[we, + we've]** seen that, / yeah. /

Table 2: Summary of features included in the final modelling, alongside actual examples from the corpus. PreC/C = Pre-core/core, PostC = post-core, Both = both Pre-core/core and post-core.

Unlike the case of pre-core/core, self-initiated self-repair (SISR) appeared in post-core positions >5% of the time, and was therefore included. A post-core has the feature **SISR** if it has either a slash unit with the tag `%` (abandoned utterance), or brackets `[]` which indicate repair in the transcriptions (Meteer & Taylor 1995). Table 2 summarises and exemplifies all the features included the final modelling.

2.3 Determination of tilt-conformity

The biases that the forms of questions impose on the answer are called *conduciveness* (Bolinger 1957, Quirk et al. 1985) or *tilt* (Heritage & C Raymond 2021). Three question design factors determine tilt: syntactic type, polarity of the question, and presence of negative polarity items.

There are three main **syntactic types** of questions: *Inverted questions* (a.k.a. *interrogative-formatted questions*) are those where the auxiliary verb precedes the subject, e.g. in *Are you eating?*, the auxiliary *are* precedes the subject *you*. *Queclaratives* (a.k.a. *declarative-formatted questions*) have the same syntax as a statement (e.g. *So you're eating.*) but serves as a question, sometimes with rising intonation. *Tag questions* consist of a declarative plus a tag that turns it into a question, usually the word *right* or an inverted auxiliary-subject sequence with polarity reversed from the statement, e.g. *You are eating, aren't you?*, where *aren't you* inverses the polarity of *you are*. The three types are largely determined from SwDA tags: inverted questions have unmodified tags, whereas queclaratives take the modifier `^d` and tag questions `^t`. Some exceptions were manually corrected; details are in Appendix B.3.

The **polarity of the question** is in most cases the polarity of the root of the question in a dependency parse: if a negator depends on it, then it is negative, otherwise it is affirmative. For tag questions, the polarity of the question is defined as the polarity of the declarative portion of the question. When a tag question has an auxiliary-subject sequence as the tag, the root is located in the tag rather than the declarative (e.g. the second *are* in *You are eating, aren't you*), so the polarity of the question is the opposite of the root. Details are in the Appendix.

Negative polarity items (NPIs) are words like *at all*, *any*, *yet* etc., which occur only in negative statements and questions, and are usually said to shift the tilt towards 'no' answers (e.g. Heritage & C Raymond 2021).

From the three question design features above, the tilts of the questions were determined following

Type	Pol	Tilt	Example
Inverted	+	yes	Are you fly fishing?
	-	yes	Isn't that correct?
Quecla-ratives	+	yes	Now this is a LeBaron?
	-	no	You can't read labels?
Tag	+	yes	Those are good aren't they?
	-	no	You don't have mountains in Texas, do you?

Table 3: Types of question syntax without NPIs and their associated tilts. Pol = polarity.

standard overviews (e.g. Heritage & Clayman 2010: 142-143, Pillet-Shore 2017, Stivers 2022: 11). Queclaratives are tilted towards the same polarity as the statement, e.g. *So you're eating?* is biased towards 'yes', *So you're not eating?* towards 'no'. Tag questions are similarly tilted towards the same polarity as the declarative portion of the question. Positive inverted questions are assumed to be biased towards 'yes' answers, e.g. *Are you eating?* is biased towards 'yes', as are negative inverted questions like *Aren't you eating?*. Table 3 summarises this situation. Questions with NPIs are assumed to be negatively-tilted, unless they are found in negative inverted questions.

Answers were sorted into tilt-conforming polarity (TC), tilt-non-conforming polarity (TNC), and no polarity (NP) by considering the polarity of the answers. Answers with cores tagged `ny`, `na`, `aa`, `sd^m` were considered positive, and those tagged `nn`, `ng`, `ar` were considered negative; these polarities were compared with the tilt of the question to determine tilt-conformity. Those tagged `arp` and `nd` (answers classified by SwDA as dispreferred) were manually annotated for polarity. Answers tagged `no`, `am` were considered NP; they are neither 'yes' nor 'no', e.g. 'maybe' or 'it depends' answers. Answers without any of these dialogue acts were excluded from the sample; they typically involve transformative answers that do not clearly give a 'yes' or 'no', but do not explicitly refuse to provide a polarity like `no`, `am` either.

2.4 Statistical analysis

The statistical approach taken is mixed mode latent class analysis (MMLCA) (Morgan 2015), which combines latent class and latent profile modelling (Nylund-Gibson & Choi 2018) by allowing both categorical and continuous variables. It identifies distinct categories of answer designs, called *latent classes*, in a data-driven way that does not predefine groups. Each latent class has a distinct distribution of feature values, as well as a prior probability

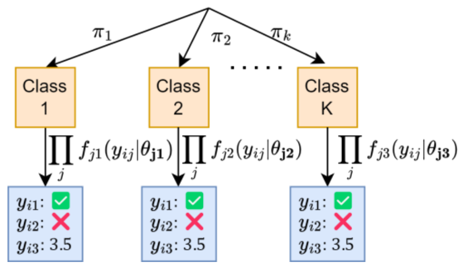

Figure 2: An illustration of the MMLCA for an answer instance with feature profile $\boldsymbol{y_i}$ = [✅, ❌, 3.5], with two dichotomous and one continuous variable.

representing how prevalent it is in the overall corpus. For each answer, the model generates the posterior probability of it belonging to each class, rather than assigning it to a single class. Examining the feature distribution of each class allows us to see and interpret answer designs holistically, abstracting over individual features.

The overall likelihood of the mixed modal latent class analysis model (MMLCA) is:

$$\prod_{i=1}^{N} f(\boldsymbol{y_i}|\boldsymbol{\Phi}) = \prod_{i=1}^{N}\left(\sum_{k=1}^{K}\pi_k\prod_{j=1}^{J} f_{jk}(y_{ij}|\boldsymbol{\theta_{jk}})\right)$$

where $\boldsymbol{y_i}$ is the profile of answer design features like fillers, discourse markers and offset time extracted for answer instance i, $\boldsymbol{\Phi}$ is the model parameters, N is sample size, K is the number of latent classes of answer designs, J is the number of features, π_k is the prior probability of an answer belonging to latent class k, and $\boldsymbol{\theta_{jk}}$ are the class-specific model parameters for the distribution of each feature j in class k. Note that the probability of the features conditional on latent class are multipled together to get their joint probability, i.e. within each latent class, features are assumed independent. For each observation, the most likely latent class is:

$$\underset{1\le k\le K}{\operatorname{argmax}}\left(\pi_k\prod_{j=1}^{J} f_{jk}(y_{ij}|\boldsymbol{\theta_{jk}})\right)$$

After fitting the model, tilt-conformity is used to predict the design of the answer with the ML three-step approach (Vermunt, 2010). The full process is implemented in MPlus (Muthén & Muthén 2019), accessed through `MPlusAutomation` in R (Hallquist & Wiley 2018).

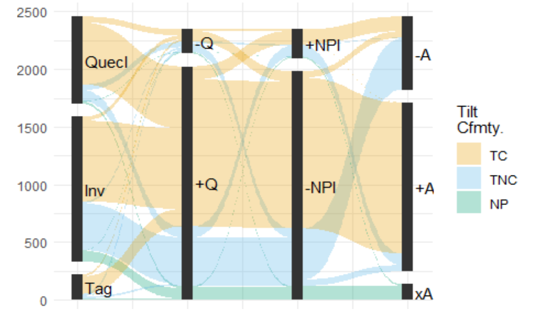

Figure 1: Sankey diagram of extracted data by tilt-related properties. Quecl = queclaratives, Inv = inverted questions, -Q and +Q = negative and affirmative questions, +NPI and -NPI = with and without NPIs, -A and +A = positive and negative answers, xA = no-polarity answers, cfmty = conformity.

3 Results

A total of $N = 2233$ Q-A pairs were extracted from the corpus, slightly more than Stivers' (2022) 1738 and considerably more than most other studies. As shown in Figure 2, there are considerable skews in tilt-related properties: Positive inverted questions without NPIs are by far the most common, followed by positive queclaratives; other categories are much rarer. Other descriptive statistics are in Appendix C; this section will focus on modelling results.

3.1 Latent classes and features

Mixed mode latent class models were run on all the binary turn design features plus OFFSET, which is modelled as Gaussians with class-varying means and variances. Models with 1-7 classes were fitted, with 8000 random starts and 4000 remaining at the final stage. Although different random starts converged to slightly different log-likelihood values, inspection of parameter estimates for top values reveals that they are almost identical.

To find the optimal number of classes, the models with 1-7 classes were compared using a variety of quantitative measures to determine the optimal model, following Nylund-Gibson & Choi (2018). This includes a series of information criteria, plus p-values of the BLRT and VLMR tests, which compare consecutive models: a significant p-value means the more complex model is better than the simpler one (Table 4). After the 5-class model, AWE shows an increase (worsening), and all other information criteria show diminishing returns clearly kicking in at the 6-class model (Figure 3). BLRT is significant for all models; VLMR is insignificant from the 4-class

#C	#Par	LL	BIC	aBIC	CAIC	AWE	BLRT	VLMR
1	15	−15,934	31983	31936	31998	32144	–	–
2	31	−14,327	28893	28794	28924	29225	<0.001	<0.001
3	47	−13,776	27915	27766	27962	28418	<0.001	<0.001
4	63	−13,472	27430	27230	27493	28105	<0.001	0.15
5	79	−13,287	27184	26933	27263	28030	<0.001	0.07
6	95	−13,176	27085	26783	27180	28103	<0.001	0.15
7	111	−13,092	27041	26688	27152	28229	<0.001	0.24

Table 4: #C = Number of classes, #Par = Number of parameters; *LL* = model log-likelihood; BIC = Bayesian information criterion; aBIC = sample size-adjusted BIC; CAIC = consistent Akaike information criterion; AWE = approximate weight of evidence criterion; BLRT = bootstrapped likelihood ratio test *p*-value; VLMR = Vuong-Lo-Mendell-Rubin adjusted likelihood ratio test *p*-value.

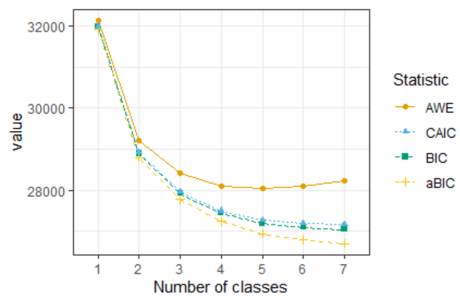

Figure 3: Information criteria for models with varying complexity. AWE worsens and aBIC, BIC and CAIC improve very little after 5 classes.

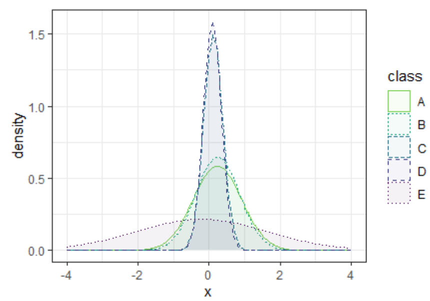

Figure 4: Model-estimated densities of offset values of the five classes. Mean offsets (in sec.) of each class are: A: .098, B: .122, C: .238, D: .249, E: -.205.

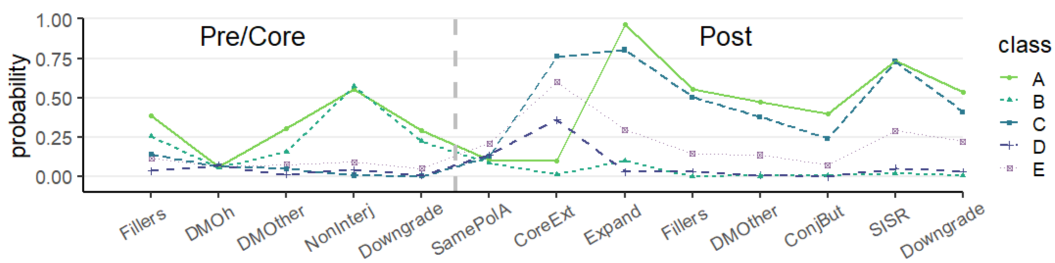

Figure 5: Estimated probabilities of each binary answer design feature by class. The fact that lines cross each other suggests that they play different functions in answer design. If all the features played similar functions and one simply uses more of them if the turn is 'more dispreferred', we would expect the lines for different classes to roughly be parallel.

Cl	Description	PreC/C fillers, DMs	Answer type	PreC/C downgrade	Core extension	Post-core expansion & fillers, DMs, etc.
A	Assumption-challenging, strongly delayed & expanded	Most	Both	Many	Very few	Most
B	Assumption-challenging, moderately delayed, unexpanded	Many	Both	Many	None	Little
C	Assumption-conforming, weakly delayed, strongly expanded	Some	Interj.	None	Most	Most
D	Assumption-conforming, undelayed & unexpanded	Few	Interj.	None	Some	Little
E	Unusual offsets	Some	Mixed	Little	Mixed	Mixed

Table 5: The five classes with key properties and brief descriptions of each class. DM = Discourse marker.

model on, though the *p*-value dipped to .07 at the 5-class model. With all metrics considered, we chose the 5-class model.

In the following paragraphs, we will answer our first research question on which answer design features tend to appear together by examining the design feature values associated with each of the five classes.

All five classes' feature profiles (Figure 5 and Figure 6) are amenable to straightforward interpretation. Sample dialogues from each class are in Appendix D. **Class A** contains strongly delayed, hedged, and lengthy answers: these are characterized by the longest offset, are often non-interjection-formatted and downgraded answers, and are most likely to have fillers and discourse markers pre-core as well as expansions and associated features like fillers and discourse markers post-core. **Class B** is like Class A, but with little post-core material and slightly less fillers and discourse markers. Inspection of transcripts also shows that they are mostly transformative, not repetitional answers. **Class C** has much shorter offsets than A-B, many fewer pre-core fillers and discourse markers, and mostly interjection-type answers, but has a similar rate of expansions as Class A. **Class D** has the shortest offsets and least pre-core material, is largely interjection-type, there are some core extensions but almost no expansion. **Class E** has greatest offset variance and largely captures instances with very long gaps or overlaps. In terms of turn design, it only stands out in having the greatest chances of SAMEPOLA, mostly due to turns with long overlaps necessitating repetition; thus, it does not shed much light on the relationship between answer design features, and will not be discussed further in the following paragraphs.

From these observations, we can group features according to the classes they are associated with. Firstly, non-interjection-type cores, pre-core/core epistemic downgrades and lack of core extensions are associated with Class A+B over C+D. These features are ASSUMPTION-CHALLENGING: They convey some stance against what is typically expected of an answer. Epistemic downgrades challenge the assumption that the answerer knows the answer with certainty. Non-interjection-type answers can reject different assumptions, e.g. challenging the relevance of the proposition raised by the questioner, assuming more control over the topics discussed, or increasing one's epistemic authority (Raymond 2003, Enfield et al. 2019, Stivers 2022); this is especially clear in the case of transformative answers, which as mentioned above are most common for Class B. The lack of core extensions is because non-interjection-type answers are already complex and thus hard to extend.

Secondly, post-core expansions and most other post-core features like downgrades, fillers, repair, discourse markers and *but* (which are most likely found in expansions rather than core extensions) are mostly associated with Class A+C over B+D. A+C may be labelled EXPANDED ANSWERS, B+D as NON-EXPANDED ANSWERS.

Finally, pre-core fillers and discourse markers follow the pattern A>B>C>D. These features DELAY the presentation of the answer core. The fact that they differ across all four classes suggests that they serve the double function of anticipating (a) assumption challenges (hence A, B > C, D) *and* (b) a longer, multi-utterance turn (hence A > B, C > D).

Interestingly, offsets pattern primarily with the first group (A, B > C, D), not other delay-related properties, as it is unclear that A>B or C>D. Thus, while our results support Kendrick & Torreira's (2015) suggestion that offset length is an aspect of turn design, silent delays may play a more restricted role than delays with fillers and particles: Longer silence primarily signals assumption-challenging answers, not expanded ones. These differences are small but noticeable: A and D are 151 ms apart.

3.2 Relationship with tilt-conformity

We now proceed to discuss how the various answer design features relate to action preference by examining their relationship with tilt-conformity, under the assumption that tilt-non-conforming answers implement dispreferred actions. Comparing

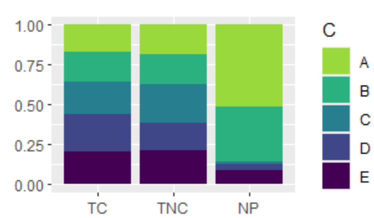

Figure 7: Distribution of probability mass assigned to each class in difference tilt-conformity conditions.

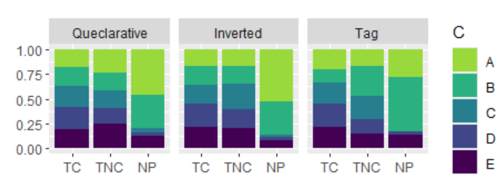

Figure 6: Distribution of probability mass assigned to each class by tilt-conformity and question type.

tilt-non-conforming (TNC) and tilt-conforming (TC) answers, D is much less probable in TNC than TC answer: the odds of getting A, B *and* C over D are higher in TNC answers (A vs D: $p = .003$; B vs D: $p = .005$; C vs D: $p < 0.001$). All other comparisons are insignificant. Comparing non-polarity-bearing (NP) answers to TC ones, the odds of A and B are significantly higher than C, D and E for NP answers ($p < 0.001$ for all); as is clear in Figure 6, TC-NP differences are much larger than TC-TNC ones, showing that assumption-challenging features are much more associated with NP than turn expansions.

To determine whether this pattern is unique to inverted questions, which dominate the sample, a by-question type barchart is given in Figure 7. The TC-TNC difference is still much smaller than TC-NP or TNC-NP. Because TNC cases are underrepresented, in most cases there is not enough power to quantitatively detect differences between TC and TNC. Visually, however, in tag questions, TNC *may* favour B (assumption-challenging, non-expanded) over not just over D ($p = .007$) but also C ($p = .105$) and A ($p = .057$), suggesting that assumption challenges play a bigger role than expansions in TNC answers to tag questions. However, a larger sample is needed to verify this.

4 Discussion and conclusion

This paper examined turn design in one context: Answers to polar questions in American English, mostly information-seeking questions due to the corpus' nature. We first examined what turn design features tend to go together. Most of the features examined fall into three categories depending on how they co-occur: assumption challenges, answer expansions, and delaying strategies. The three typical sets of strategies traditionally said to characterise dispreferred turn formats (Pillet-Shore 2017) – qualification, accounts (i.e. answer explanations) and delays – fall into these three categories. This suggests that the three types of strategies have distinct distributions and thus functions.

Two unexpected observations emerge from this typology. Firstly, while the choice between interjection- vs. non-interjection-type answers is usually associated with a separate dimension (G Raymond 2003) from the dispreferred turn design strategies of qualification, account and delay, we find that it patterns with qualification in the assumption-challenging category. Indeed, only 5% of interjection-type answers are downgraded, while 21% of non-interjection-type answers are. Secondly, offset patterns with assumption-challenging features rather than other (nonsilent) delay-related features, suggesting that silent delays project only assumption-challenging, not expanded answers.

The fact that nonsilent delays correlate with both assumption challenges and answer expansions may be explained by multiple mechanisms. Firstly, they may anticipate the other turn design features, e.g. Heritage (2015) argues that *well* alerts the listener to upcoming nonstraightforward, transformative *and* expanded answers. They may also directly signal similar meanings as some other answer design strategies, e.g. difficulty in memory retrieval or lower level of knowledge (Smith & Clark 1993, Brennan & Williams 1995), which presumably correlate with epistemic downgrades.

To examine how action preference is related to answer design, we also examined the relationship between tilt-conformity and answer design. As expected, tilt-nonconformity disfavours answers with no delays, expansions, *or* assumption-challenging features over answers with at least some of these. TNC status may favour assumption-challenging features even more in tag questions, probably because they have stronger tilts, and thus going against the tilt poses a greater face threat. Yet, regardless of question type, the tilt-conformity effect is far smaller than the difference between non-polarity-conveying and polarity-conveying answers (regardless of tilt-conformity): Answers without polarity are overwhelmingly designed with non-interjection-type answers and/or epistemic downgrades, likely because they inherently challenge the assumption that the answerer is willing and able to give a straightforward yes/no. This extends Robinson's (2020a) hypothesis that 'yes' and 'no' answers are both preferred answers to positive inverted questions, and only conditional answers are dispreferred, by expanding it to all polar question formats with non-polarity-bearing answers. One difference between Robinson's and our study is that he found no significant difference in pre-beginning behaviour (including fillers and discourse markers in our study) between tilt-conforming and tilt-nonconforming answers, while we do find that tilt-nonconforming answers disfavour class D, which has the least pre-beginning behaviour. This is likely a result of our larger sample size, and supports Robinson's idea that although the *social action* of asking a positive inverted question doesn't by itself impose a preference, the syntactic form still encodes a tilt (Robinson 2020b).

Our results favour a gradient, multidimensional view of preference (Robinson 2020a). Limited by the categories employed by pre-existing SwDA annotations, our study cannot fully examine this richness, e.g. we could not distinguish between expansion types or determine which questions are truly information-seeking. Future studies will hopefully shed further light on these dimensions, a key piece of research as dialogue systems strive to mimic human conversational behaviour (Alloatti et al. 2021, Dingemanse & Liesenfeld 2022, Lah & Lee 2023).

References

Alloatti, Francesca, Luigi Di Caro, Alessio Bosca, & others. 2021. Conversation analysis, repair sequences and human computer interaction–a theoretical framework and an empirical proposal of action. In *Proceedings of the Fourth Workshop on Reasoning and Learning for Human-Machine Dialogues (DEEP-DIAL 2021) at the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)*. Association for the Advancement of Artificial Intelligence.

Asparouhov, Tihomir & Bengt Muthén. 2014. Auxiliary variables in mixture modeling: Three-step approaches using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal* 21(3). 329–341.

Bolinger, Dwight. 1957. *Interrogative structures of American English: the direct question* (Publication of the American Dialect Society; 28; No. 28). Alabama: American Dialect Society.

Brennan, Susan E & Maurice Williams. 1995. The feeling of another′ s knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language* 34(3). 383–398.

Calhoun, Sasha, Jean Carletta, Jason M Brenier, Neil Mayo, Dan Jurafsky, Mark Steedman & David Beaver. 2010. The NXT-format Switchboard Corpus: a rich resource for investigating the syntax, semantics, pragmatics and prosody of dialogue. *Language Resources and Evaluation* 44. 387–419.

Cappelli, Gloria. 2007. *"I reckon I know how Leonardo da Vinci must have felt ...": epistemicity, evidentiality and English verbs of cognitive attitude*. Pari (Grosseto): Pari Publishing.

Dingemanse, Mark & Andreas Liesenfeld. 2022. From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 5614–5633. Dublin, Ireland: Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-long.385.

Enfield, N. J., Tanya Stivers, Penelope Brown, Christina Englert, Katariina Harjunpää, Makoto Hayashi, Trine Heinemann, et al. 2019. Polar answers. *Journal of Linguistics* 55(2). 277–304. https://doi.org/10.1017/S0022226718000336.

Fischer, Kerstin. 2015. Conversation, Construction Grammar, and cognition. *Language and Cognition* 7(4). 563–588. https://doi.org/10.1017/langcog.2015.23.

Ford, Cecilia E. 2001. At the intersection of turn and sequence. In Margaret Selting & Elzabeth Couper-Kuhlen (eds.), *Studies in interactional linguistics* (*Studies in Discourse and Grammar* 10). Amsterdam: John Benjamins Publishing. 51–80.

Jurafsky, Daniel, Elizabeth Shriberg & Debra Biasca. 1997. Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual.

Godfrey, John J, Edward C Holliman & Jane McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. In *IEEE International Vonference on Acoustics, Dpeech, and Signal Processing*, vol. 1, 517–520. IEEE Computer Society.

Hallquist, Michael N. & Joshua F. Wiley. 2018. *MplusAutomation* : An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal* 25(4). 621–638. https://doi.org/10.1080/10705511.2017.1402334.

Heritage, John. 1998. *Oh*-prefaced responses to inquiry. *Language in Society* 27(3). 291–334. https://doi.org/10.1017/S0047404500019990.

Heritage, John. 2005. Cognition in discourse. In Hedwig Te Molder & Jonathan Potter (eds.), *Conversation and Cognition*, 184–202. 1st edn. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511489990.009.

Heritage, John. 2015. *Well*-prefaced turns in English conversation: A conversation analytic perspective. *Journal of Pragmatics* 88. 88–104.

Heritage, John & Steven Clayman. 2010. *Talk in action: interactions, identities, and institutions* (Language in Society 38). Chichester Malden: Wiley-Blackwell.

Heritage, John & Geoffrey Raymond. 2012. Navigating epistemic landscapes: Acquiescence, agency and resistance in responses to polar questions. In Jan P. De Ruiter (ed.), *Questions*, 179–192. 1st edn. Cambridge University Press. https://doi.org/10.1017/CBO9781139045414.013.

Heritage, John & Chase Wesley Raymond. 2021. Preference and polarity: epistemic stance in question design. *Research on Language and Social Interaction* 54(1). 39–59. https://doi.org/10.1080/08351813.2020.1864155.

Honnibal, Matthew & Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

Kendrick, Kobin H. & Francisco Torreira. 2015. The timing and construction of preference: a quantitative study. *Discourse Processes* 52(4). 255–289. https://doi.org/10.1080/0163853X.2014.955997.

Kendrick, Kobin H & Judith Holler. 2017. Gaze direction signals response preference in conversation. *Research on Language and Social Interaction* 50(1). 12–32.

Lah, Ji Young & Yo-An Lee. 2023. Managing turn-taking through beep sounds by Bixby: Applying conversation analysis to human-chatbot interaction constructions. *Linguistic Research* 40. 61–87. https://doi.org/10.17250/KHISLI.40..202309.003.

Lee, Seung-Hee. 2015. Two forms of affirmative responses to polar questions. *Discourse Processes* 52(1). 21–46. https://doi.org/10.1080/0163853X.2014.899001.

Masyn, Katherine E. 2013. Latent Class Analysis and Finite Mixture Modeling. In Todd D. Litte (ed.), *The Oxford Handbook of Qauntitative Methods in Psychology: Vol 2: Statistical Analysis*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199934898.013.0025.

Meteer, Marie & Ann Taylor. 1995. *Dysfluency Annotation Stylebook for the Switchboard Corpus*.

Muthén, Bengt & Linda Muthén. 2019. Mplus: A general latent variable modeling program.

Morgan, Grant B. 2015. Mixed mode latent class analysis: An examination of fit index performance for classification. *Structural Equation Modeling: A Multidisciplinary Journal* 22(1). 76–86. https://doi.org/10.1080/10705511.2014.935751.

Nishizaka, Aug & Kaoru Hayano. 2015. Conversational Preference. In Karen Tracy, Todd Sandel & Cornelia Ilie (eds.), *The International Encyclopedia of Language and Social Interaction*, 1–7. 1st edn. Wiley. https://doi.org/10.1002/9781118611463.wbielsi071.

Nylund-Gibson, Karen & Andrew Young Choi. 2018. Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science* 4(4). 440.

Pillet-Shore, Danielle. 2017. Preference Organization. In *Oxford Research Encyclopedia of Communication*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190228613.013.132.

Pomerantz, Anita. 1985. Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes. In J. Maxwell Atkinson (ed.), *Structures of Social Action*, 57–101. 1st edn. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511665868.008.

Pomerantz, Anita & John Heritage. 2012. Preference. In Jack Sidnell & Tanya Stivers (eds.), *The Handbook of Conversation Analysis*, 210–228. 1st edn. Wiley. https://doi.org/10.1002/9781118325001.ch11.

Potts, Christopher. 2011. The Switchboard Dialog Act Corpus. Computational Pragmatics. http://compprag.christopherpotts.net/swda.html.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Raymond, Geoffrey. 2003. Grammar and social organization: yes/no interrogatives and the structure of responding. *American Sociological Review* 68(6). 939–967. https://doi.org/10.1177/000312240306800607.

Robinson, Jeffrey D. 2020a. Revisiting preference organization in context: a qualitative and quantitative examination of responses to information seeking. *Research on Language and Social Interaction* 53(2). 197–222. https://doi.org/10.1080/08351813.2020.1739398.

Robinson, Jeffrey D. 2020b. One Type of Polar, Information-Seeking Question and Its Stance of Probability: Implications for the Preference for Agreement. *Research on Language and Social Interaction* 53(4). 425–442. https://doi.org/10.1080/08351813.2020.1826759.

Roberts, Seán G, Francisco Torreira & Stephen C Levinson. 2015. The effects of processing and sequence organization on the timing of turn taking: a corpus study. *Frontiers in psychology* SA 2015(6). 509.

Sacks, Harvey. 2010 [1987]. On the preferences for agreement and contiguity in sequences in conversation. In Patrick Griffiths, Andrew Merrison & Aileen Bloomer (eds.), *Language in Use: A Reader*, 8–22. Milton Park: Routledge.

Smith, Vicki L & Herbert H Clark. 1993. On the course of answering questions. *Journal of Memory and Language* 32(1). 25–38.

Stivers, Tanya, Nicholas J. Enfield, Penelope Brown, Christina Englert, Makoto Hayashi, Trine Heinemann, Gertie Hoymann, Federico Rossano, Jan Peter De Ruiter & Kyung-Eun Yoon. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* 106(26). 10587–10592.

Stivers, Tanya & Makoto Hayashi. 2010. Transformative answers: One way to resist a question's constraints. *Language in Society* 39(1). 1–25. https://doi.org/10.1017/S0047404509990637.

Stivers, Tanya. 2022. *The book of answers: alignment, autonomy, and affiliation in social interaction* (Foundations of Human Interaction). New York, NY: Oxford University Press.

Thompson, Sandra A. 2002. "Object complements" and conversation: towards a realistic account. *Studies in Language* 26(1). 125–163. https://doi.org/10.1075/sl.26.1.05tho.

Vermunt, Jeroen K. 2010. Latent class modeling with covariates: Two improved three-step approaches. *Political analysis* 18(4). 450–469.

Appendices

A Switchboard tags

`qy`	polar question
`ny`	'yes' answer
`nn`	'no' answer
`ny`	affirmative non-'yes' answer
`ng`	negative non-'no' answer
`no`	other answer
`nd`	dispreferred answer
`aa`	acceptance
`aap`	partial acceptance
`am`	'maybe' answer
`ar`	rejection
`arp`	partial rejection
`h`	hold
`^r`	self-repetition
`^m`	other-repetition
`^e`	expansion
`^g`	tag question
`^h`	hedge
`sd`	statement, not opinion
`sv`	statement, opinion
`{F }`	filler
`{D }`	discourse marker
`{C }`	conjunction
`%`	abandoned utterance
`[]`	repair
`<>`	vocalism

B Details of feature extraction

B.1 Details of extracting question-answer pairs and answer features

Before further processing, any slash unit with + as its dialogue act was merged with the preceding act by the same participant. When there are additional slash units after the first question slash unit of a certain turn (for example, reformulations of the question or turn increments), all slash units up to either the the slash unit right before the start of the next turn or the one right after the start of the next turn were considered, whichever one's midpoint was closer to the start of the next turn.

The last question slash unit of the question turn was considered in determing question type and polarity. This question was parsed with spaCy. If spaCy identified multiple sentences within the slash unit, then we took the one with a question mark if there is only one such slash unit; we took the longest sentence with a question mark if there were multiple such slash units; and we took the longest sentence if there were no question marks.

The following were treated as potential answer cores: `ny` (yes answers), `nn` (no answers), `na` (affirmative non-yes answers), `ng` (negative non-no answers), `no` ('other answers'), `sd^m` (repetition of the other's question, which generally affirm the answer in this corpus), `aa` and `ar` (acceptance / rejection of question-formatted collaborate completions), plus any `sd` with the word 'depend' in it. For each responsive turn, the first slash unit with one of these dialogue acts was treated as core. Some yes/no answers were mistakenly tagged as `b` (backchannels); when they are classified as interjection-type answers (see below) and there are no other slash units in the response, they are treated as 'yes' answers. Although `sv` and `sd` often also implemented polar answers, they were not included as it is difficult to automatically determine whether they bear polarity and, if so, whether they are positive or negative. Determination of answer polarity was discussed in the main text.

Well and *you know* were originally extracted separately from other discourse markers, but later merged into the general category.

OFFSET, SISR and NONINTERJ were mostly extracted as stated in the main text; NONINTERJ are those answers classed as `nn` and `ny`. In addition, a small number of answers from other classes were also interjection-type. These were extracted by considering a list of potential interjection-type answers: *yeah*, *no*, *yes*, *uh-huh*, *right*, *huh-uh*, *okay*, *sure*, *exactly*, *absolutely*, *definitely*, *certainly*, *probably*, *yep*, *yip*, *mm-hm*, *of course*, *no question*, *I'll say*, *possibly*, *maybe*, *alright*, *fine*. This list combines the one in Stivers (2022), plus other interjection-type answers fouund in an inspection of all one-word cores attested in the corpus. An answer is considered interjection-type if its core contains one of these interjections alone, or one of these interjections after by *uh, um, oh, well*.

The determination of DOWNGRADE was relatively complex. Lists of adverbs and auxiliaries were created by parsing all the answer (pre-)cores in the corpus, extracting all adverbs and auxiliaries, and determining polarity. Auxiliaries deemed to be downgraders include *could*, *might*, *should*, *may*, *can*, *ought*, *must*. Adverbs deemed to be downgraders on their own were *probably*, *somewhat*, *sometimes*, *personally*, *maybe*, *perhaps*, *possibly*, *fairly*. Adverbs deemed to be downgraders when combined with negation were *really*, *so*, *very*, *too*, *usually*, *exactly, normally*, *particularly, always*; these were only considered downgraders when there is a negator in the same sentence.

Epistemic verbs include the lemmas *think*, *believe*, *guess*, *suppose*, *know*, *feel*, *hear*, *assume*, *bet*, *conjecture*, *consider*, *doubt*, *expect*, *fancy*, *figure*, *reckon*, *gather*, *imagine*, *judge*, *presume*, *sense*, *surmise*, *suspect*, *trust* with *I* as subject, and *say* with subjects other than *I*. Other phrases included were *my guess*, *my feeling*, *I get the feeling*, *looks like.*

B.2 Unused answer design features

The following features were extracted but not used in the end because they appeared less than 5% of the time.

A pre-core/core has the feature HOLD if it contains a slash unit tagged `h` (hold).

Non-linguistic vocalisms are transcribed in the corpus within angular brackets <>. Four were coded into features: Throat-clearing (THROAT) from the tag `<throat_clearing>`, laughter (LAUGH) from the tag `<laughter>`, lip-smacking (LIPSM) from the tag `<lipsmack>`, and breaths (BREATH) from the tag `<breathing>`.

Conjunctions (CONJ) marked `{C }`, with the forms *so*, *but*, *because*, and sentence-initial *And* were treated as conjunctions. Edit terms (EDITTERM) were extracted with `{E }`, with *I mean* originally extracted apart from other edit terms; all edit terms were discarded in the end.

The feature DIFFPOLA was used for dialogue acts conveying a different polarity as the core.

Sure, *exactly* and *really* were considered UPGRADER when not accompanied by negators. *Absolutely*, *definitely* and *certainly* were always considered upgraders.

B.3 Determination of tilt-conformity

Generally, any question without an auxiliary-subject (or copula-subject) sequence or a tag is considered queclarative. This include subclausal questions. The main exception is that when a question omits a copula or auxiliary verb that cannot be omitted in declaratives; in this case, this is considered ellipsis of the beginning of the question (Quirk et al. 1985), e.g. *you got any hobbies that you want to talk about?*. For questions starting with *how about* (e.g. *{C And } how about SILENCE OF THE LAMB? /*), the question type was set to be the same as that of the previous question.

In general, question slash units with `^d` were treated as queclaratives, those with `^g` as tag questions, and other questions were treated as inverted. Sub-clausal questions were treated as declarative. However, there are a number of cases where the Switchboard corpus appeared to use intonation instead of syntax to determine `^d` would be used. To smooth out these inconsistencies, if a question was tagged as inverted but our syntactic parse finds an auxiliary-subject sequence, or the other way around, we manually checked them to determine question type.

Polarity was determined as described in the main text: For all questions but tag questions with auxiliary-subject tags, it was whether the root had a negator dependent; for tags with auxiliary-subject tags, it was the opposite polarity as the tag.

Answer polarity largely was determined as mentioned in the main text. Answers tagged `sd` containing the word *depend* were treated as NP.

C Descriptive statistics

In the main text, we have discussed the model results. In this appendix we present the descriptive statistics to paint a more comprehensive picture of the data.

Relationships among binary turn design features. To examine the relationship between different binary variables, log-odds ratios were computed between each pair of features, and plotted in Figure 8. Positive values mean the features tend to appear together, negative ones mean they tend to appear apart, and zero means no relationship. As is clear from the heatmap, most relationships are non-negative. Most strong positive relationships are concentrated between features of the post-core and, to a lesser extent, between features of the core/pre-core. EXPAND and post-core SISR are especially notable for their strong association with other post-core features, suggesting most of those other features are found in expansions. DMOH, COREEXT and SAMEPOLA are weakly or negatively correlated with other variables, and appear to work independently of other features.

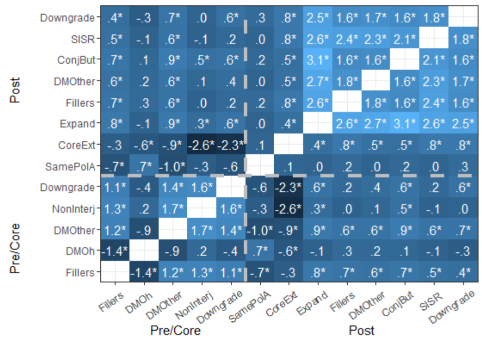

Figure 8: Log-odds ratios between different answer design features. * indicates that the two variables are significantly associated at the .05 level of significance using Fisher's exact tests.

Relationship between OFFSET *and binary turn design features*. Following Kendrick & Torreira (2015), we examine at entire distributions of offsets rather than just means. For each turn design feature, kernel density estimates of the offset were calculated when the feature is present vs when it is absent. The difference between the two densities at various values on (-2, 2) is shown in Figure 9. The clearest pattern is that for all turn design features but DMOH and EXPAND, near-zero (i.e. no gap, no overlap) onsets are much more common when the feature is absent than when it is present. However, the prevalence of gaps over overlaps only seems to be associated with the presence of the pre-core FILLERS and post-core SISR, CONJBUT, and DMOTHER features. For DOWNGRADE and NONINTERJ, longer gaps are associated with the presence of the feature, but so are

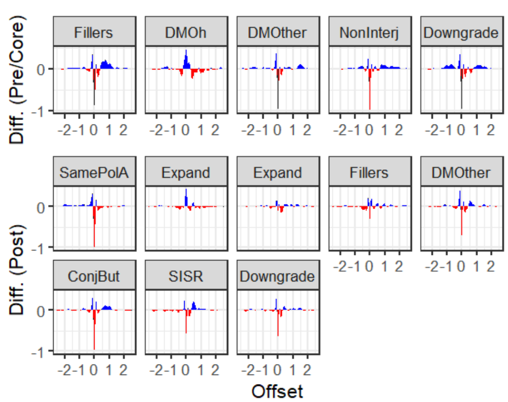

Figure 9: Difference in kernel density estimates of the OFFSET feature when each feature is present vs absent. Red (<0) means that offset value is more common when the feature is absent is larger, and vice versa.

slight overlaps; only short gaps are associated with absence. For most other features, the pattern is unclear, or even reversed for SAMEPOLA.

Relationship between tilt-conformity and binary turn design features. Generally, tilt-non-conforming (TNC) turns are more likely to contain the turn design features examined than tilt-conforming (TC) ones, and no-polarity (NP) answers are more likely to contain them than TNC ones, though the degree varies. For pre-core/core NONINTERJ, DMOTHER and DOWNGRADE, the TC-TNC difference is much smaller than the NP-TNC difference; for pre-core FILLERS or post-core STNONEXPAND, the TNC-TC difference and NP-TNC difference are more comparable. DMOH, EXPAND and SAMEPOLA are again exceptions to the general pattern.

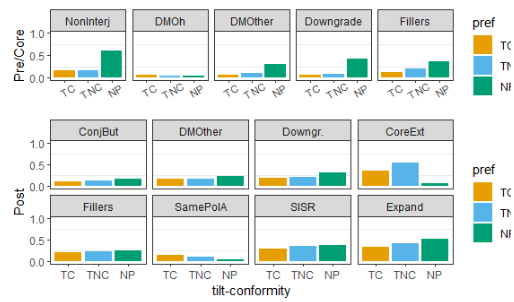

Figure 10: Barcharts of the prevalence of design features in each condition

Relationship between tilt-conformity and offsets. Near-0 offsets are most commonly seen with TC answers, followed by TNC, and finally NP. Gaps between .3-.6 seconds are most likely TNC, followed by NP and TC; beyond around .8 seconds, the order is NP > TNC > TC. From all this, it is clear that NP responses are most closely associated with long gaps, followed by TNC and TC. Nevertheless, the differences are quite minute.

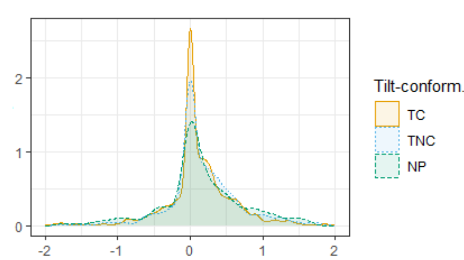

Figure 11: Kernel density of offsets by tilt-conformity.

Zeroing in on inverted questions, we find that positive inverted questions follow the general pattern in Figure 11, but negative questions are radically different: TC (positive) answers actually are *more* likely to have long gaps than TNC (negative) or NP ones (Figure 12). This may be because negative inverted interrogatives still express the speaker's stance that something in the context makes the state of affairs expressed in the question improbable (Heritage & C Raymond 2021).

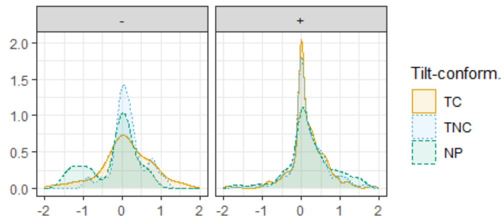

Figure 12: Kernel density estimates of offsets by tilt-conformity for inverted questions without NPIs only.

D Sample answers from the four classes

All answers given in this section have class probability of at least .95.

D.1 Class A

B Is Pennsylvania kind of out of line there? /
A {D Well, } {D actually, } I don't think they're out of line. /
[De-, + Devil's] advocate possibly, /
{C but } <rustling> <inhaling> [it, + you] are trying to avoid paying taxes /
{C and } [whe-, + whether] or not you agree with that law, [i-, + you're] still circumventing it. /
You are legal [in, + in] your circumvention of that law. /

Delays: YES – long.
Fillers and discourse markers: *Well, actually.*
Epistemic downgrade: *I don't think.*
Non-interjection answer: Repetitional, not a direct *no.*
Expansion: extensive justification and elaboration after core.

A [You don't, + {F uh, } you're not] [in-, + into] hacking or whatever <laughter>. /
B {F Oh, } [[I, + I think I'm,] + I think I'm] a hacker, /
{C but } I'm [[not, + not kind,] + not [the, + {F uh, } the,]] {D you know, } dial around randomly trying to break into computers type -- -- hackers, /
no, /
that's <laughter> one of those sports I don't go for. /

Delays: YES – long.
Fillers and discourse markers: *Oh, uh, you know.*
Non-interjection answer: Repetition, not a direct *yes*.
Epistemic downgrade: *I think.*
Expansion: extensive justification and elaboration after core.

A {D Well } [don't most of them, + doesn't just] about everything now have both metric and English. /
B They do, /
{C but } things are generally packaged in the English sized packages, {D you know. } /
You buy a quart of milk, /
{C and } sure it [has, + has] the metric equivalent written on there, /
{C but } it still a quart. /

Delays: YES – long.
Fillers and discourse markers: *you know.*
Non-interjection answer: Repetitional, not a direct *yes.*
Expansion: extensive justification and elaboration after core.

D.2 Class B

B Do you have any children? /
A {F Uh, } they're all grown up. /

Delays: YES – moderate.
Fillers and discourse markers: *Uh.*
Non-interjection answer: Transformative, not a direct *no.*
Expansion: NONE, no elaboration or justification after the non-interjection answer.

B Have you read that? /
A {F Uh, } I haven't gotten through <laughter> it yet. /

Delays: YES – moderate.
Fillers and discourse markers: *Uh.*
Non-interjection answer: Transformative, not a direct *no.*
Expansion: NONE, no elaboration or justification after the non-interjection answer.

B Did you all ever watch that? /
A [I, + {D yeah, } I] started, too, and, {F uh } -- -- [kind of, + kind of] worked away from that. /

Delays: YES – moderate.
Fillers and discourse markers: *yeah, uh.*
Epistemic downgrade: *kind of.*

Non-interjection answer: Transformative, not a direct *no*.
Expansion: NONE, no elaboration or justification after the non-interjection answer.

D.3 Class C

A Do you find trouble keeping the records for taxes and all that /
B No, /
it's not hard, /
I just keep it in a notebook and write down what I've made and, {F uh, } {D you know, } what it's going to have to go for that month /
{C and } -- -- {D you know, } it's [not that, + not that] hard. Not at all. /

Delays: YES – minimal.
Fillers and discourse markers: NONE before core.
Interjection-type answer: *No*
Expansion: significant elaboration after the core.

A [Have you, + have you] ever done anything at all? /
B Yeah, /
I have. /
{F Uh, } sit-ups /
{C or, } [al-, + also] last summer I was doing Nautilus /
{C or } last year <cough> I'm, {F uh, } belong to a club right here. /
Got kind of expensive, {F uh, } [to r-, + to [r-, + renew.]] They wanted another fifty dollars. /

Delays: NONE.
Fillers and discourse markers: NONE before core.
Interjection-type answer: *Yeah.*
Expansion: significant elaboration after the core.

A I wonder if she's written anything really recently, if she's got anything [printed, + in print.] /
B Yeah, /
she has, /
{C because } [I, + I] remember seeing a new book by her -- -- that was out, /
{C and } I think [it was a, + it was an] adult book. /

Delays: NONE.
Fillers and discourse markers: NONE.
Interjection-type answer: *Yeah.*
Expansion: significant elaboration after the core.

D.4 Class D

B When you did your papering did you start in the middle of the wall? /
A No /
I didn't. /

Delays: NONE.
Interjection-type answer: *No.*
Expansion: NONE, only extension *I didn't.*

A Have you ever read anything by Susan Howatch? /
B Yes, /
I have. //

Delays: NONE.
Interjection-type answer: *Yes.*
Expansion: NONE, only extension *I have*.

A Like, Queen's Reich, if you ever heard of them. /
B {F Oh, } sure. /
Of course. /

Delays: NONE.
Fillers and discourse markers: *Oh*
Interjection-type answer: *sure.*
Expansion: NONE, only extension *of course*.

D.5 Class E

A {C so. } [Have you, + do you] have a computer for yourself at home? /
B [Offset = 1.21] No /
I didn't. /

Delays: YES – long.
Fillers and discourse markers: NONE.
Interjection-type answer: *No.*
Expansion: NONE, only extension *I didn't.*

B [Do you work with, + do you work around] children when you work? /
A [Offset = -.70] No, /
no, /
not at all. /
I work with <noise> computers. /

Delays: NONE – overlap of speakers.
Fillers and discourse markers: NONE.
Interjection-type answer: *No.*
Expansion: elaboration after the core.

A Do you have any [l-, +] nieces or nephews
<Laughter> ((then)) ? /
B [Offset = -2.09] Yeah. /
Yeah. /
I have a nephew. /
He's a little brat. /

Delays: NONE – overlap of speakers.
Fillers and discourse markers: NONE.
Interjection-type answer: *Yeah.*
Expansion: elaboration after the core