Article

Noa Attali, Amnon Attali and Anne Therese Frederiksen*

When one wrong rights another: speakers passivize to express the subject as the experiencer in psychological verb use

https://doi.org/10.1515/cllt-2024-0091
Received August 30, 2024; accepted February 20, 2025; published online April 10, 2025

Abstract: Speakers have many choices for how to express a given message yet strongly prefer some over others. We suggest that one preference speakers have in spontaneous language use is presenting the experiencer of an emotion in a salient position: as the subject of the sentence. We extend prior work on the experiencer-as-subject bias to the understudied area of spontaneous language use. We examined a corpus for speakers' choices regarding a) active versus passive sentences and b) two kinds of psychological verbs, subject-experiencer versus object-experiencer. Language users generally prefer active over passive sentences, and there is some evidence that object-experiencer verbs (e.g., Mary frightens Lisa) are dispreferred compared to subject-experiencer verbs (e.g., Lisa fears Mary) much as passive sentences are compared to active sentences. However, in line with an experiencer-as-subject bias, we show how one wrong can right another: speakers show a greater-than-expected preference to produce a generally dispreferred sentence structure, when doing so achieves the desired experiencer-as-subject. Specifically, we find stark differences in passive use rates: as many as 50 % of the uses of certain verbs in the dispreferred object-experiencer verb class are passive, while 4 % is the highest rate of passive sentences for verbs in the preferred subject-experiencer verb class.

Keywords: psychological verbs; experiencer verbs; spontaneous language production; passives; argument structure

***Corresponding author: Anne Therese Frederiksen**, Department of Communication Arts, Sciences, and Disorders, Brooklyn College, CUNY, Brooklyn, NY, USA, E-mail: anne.frederiksen@brooklyn.cuny.edu
Noa Attali, Department of Language Science, University of California Irvine, Irvine, CA, USA
Amnon Attali, Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Figure 1: A scenario which is truthfully described by the utterances in (1)-(4).

1 Introduction

Language users have many different ways to express whatever message they have in mind in everyday conversation. How do people choose between alternative expressions of a message? For example, how would people choose between the sentences (1)-(4), which all truthfully describe the scenario in Figure 1? Among the many potential motivating factors, we focus here on whether speakers prefer to mention the experiencer of an emotion as the sentence subject in naturalistic speech. Previous studies have shown that sentences are easier to understand when the experiencer of an emotion is also the sentence subject, but it has not yet been shown whether people indeed prefer to produce experiencers as subjects in spontaneous language use. Here, we conducted a large-scale corpus study to investigate this question.

As an illustration of the potential experiencer-as-subject bias, consider the choice between the alternatives (1)-(4). (In these examples, we bold **the experiencer of the emotion** and underline the subject of the sentence.) For the scenario in Figure 1, and in all four of its descriptions (1)-(4), *Mary* is the experiencer of the fear emotion while *Lisa* is the cause of the emotion. But the experiencer *Mary* is the subject only in sentences (1) and (4). In sentences (2) and (3), the stimulus *Lisa* is the subject. So only sentences (1) and (4) achieve the desired property of having the experiencer as the subject.

(1) **Mary** fears Lisa.

(2) Lisa is feared by **Mary**.

(3) Lisa frightens **Mary**.

(4) **Mary** is frightened by Lisa.

However, the two sentences achieve this experiencer-as-subject goal in different ways: the sentence in (1) uses active voice (as does (3)) while the sentence in (4) uses passive voice (as does (2)). This is no trivial difference: passive voice is much less frequent and harder to process than active voice (Dick and Elman 2001; Ferreira 2003; Nguyen 2021), which might mean that (4) should be generally dispreferred. But we highlight in this study that, when using the passive achieves another goal, speakers may be more willing to use it. Specifically, for certain verbs, the passive voice enables the experiencer to be the subject. Experiencer-as-subject is already achieved with the active voice (1) for a verb like *fear*, since the active voice allows the experiencer to be presented in the subject position. But for a verb like *frighten*, the *passive* (4) rather than the active (3) alternative places the experiencer as the subject.

Thus, a broad prediction emerges of a higher rate of passive sentences for verbs like *frighten* than for verbs like *fear*. Specifically, the preference to say (4) rather than (3) should be greater than the preference to say (2) rather than (1) (Equation (1)). Several previous studies of elicited spoken English and written English have found just this difference in passive rates (Do and Kaiser 2022; Ferreira 1994; Gennari and MacDonald 2009), but whether the prediction also holds for spontaneous speech is unclear.

$$\frac{\text{PASSIVE USES OF FRIGHTEN-LIKE VERBS (4)}}{\text{TOTAL FRIGHTEN-LIKE VERB USES ((4) + (3))}} > \frac{\text{PASSIVE USES OF FEAR-LIKE VERBS (2)}}{\text{TOTAL FEAR-LIKE VERB USES ((2) + (1))}} \quad (1)$$

To preview our results, we confirm based on corpus data of spontaneous speech that speakers show a greater preference for (4) as compared with (3), relative to their dispreference for (2) as compared with (1). In general, we find that verbs like *frighten* are sometimes passivized as often as half the time, whereas verbs like *fear* are always rarely passivized. This stark contrast in passive sentence rates suggests that one wrong can help to right another: the dispreferred passive can be selected when its use avoids another dispreferred situation, that of not mentioning the experiencer as the subject.

1.1 Alignment in passive sentences

Although attested across languages (Grohmann et al. 2021; Keenan and Dryer 2007), the passive voice is infrequent. Corpora of English child-directed speech show a rate of passives around 0.5 % (Gordon and Chafetz 1990; Nguyen 2021), suggesting that parents avoid using the passive when talking to their children. Corpora of adult-to-adult speech show ratios of active to passive sentences between 35:1 and 10:1, and although the ratio increases to ~6:1 in written language (and further depends on

factors such as discourse register), passives are still much less frequent than actives overall (Dick and Elman 2001; Roland et al. 2007). Similarly, passive voice is more difficult to process than active voice. For example, listeners are more likely to misinterpret a passive rather than an active sentence, especially when the meaning of the sentence is implausible (Ferreira 2003). Similarly, Gibson et al. (2013) found that people are more likely to rely on plausibility information to interpret passive rather than active structures.

What makes the passive dispreferred? One argument focuses on its lower frequency of use as a source of difficulty in itself: for example, that distributional frequencies affect how children learn linguistic structures (e.g., Plunkett and Marchman 1993), or more broadly that experience of the statistical patterns of language plays an important role in natural language processing (e.g., Wells et al. 2009). However, the argument that we focus on here is concerned with another difference between the passive and active voice: how they align elements of meaning and structure. Consider the active sentence (5): on the meaning level (in particular, the level of thematic roles), *Mary* (bolded) is the thematic agent of the kicking and *the ball* is the thematic patient.[1] On the structural level, *Mary* (underlined) is the subject and *the ball* is the object of the verb. So, the active voice sentence (5) maps the agent to the subject and the patient to the object. The passive voice sentence (6) displays the opposite mapping. Here, *Mary* is still the agent and *the ball* remains the patient, but *the ball*, not *Mary*, is now mapped to the subject.

(5) **Mary** kicked the ball.

(6) The ball was kicked by **Mary**.

Early theoretical accounts posited a hierarchical relationship between thematic roles and argued that speakers must align the thematic role highest in the hierarchy with the "highest" or most prominent syntactic role, the sentence subject (e.g., Baker 1988; Belletti and Rizzi 1988). The agent role (or the 'agent-like' role, in a prototype rather than categorical notion of argument) is highest in the hierarchy and should be aligned with the subject whenever it is present (e.g., Baker 1988, 1997; Dowty 1991). However, psycholinguistic accounts treat alignment as a processing preference (e.g., as the result of a processing strategy that tries to generate less complex structures; see Bornkessel and Schlesewsky 2006). Still other accounts have argued for a similar outcome but have linked it to a preference for producing the more accessible or prominent words and phrases (including those mapped to prominent thematic roles) early in the sentence (Bock 1982; Ferreira 1994). The difference in mapping between sentence position and thematic roles is then termed a difference in *alignment*. In

1 Also referred to as the theme.

other words, the active voice (5) is aligned, while the passive voice (6) is misaligned. In general, the concept of alignment extends beyond thematic roles and sentence position: for example, Wagers and Pendleton (2016) illustrate how animate but not inanimate arguments predict subject gaps, facilitating subject-relative clause processing.

Why misaligned mappings should be dispreferred is still an open question. One approach, highlighting language experience again, says that it is a useful heuristic for language processing to assume common mappings. Regarding our examples (5) and (6) – as examples of agent-patient verbs in English – it would be useful to assume that these subject-verb-object sentences should be interpreted as agent-verb-patient, leading to a general agent-as-subject mapping preference. A model of parsing called the Late Assignment of Syntax Theory (LAST; Townsend and Bever 2001) argues that sentence processors first compute a "quick and dirty" interpretation of sentences which is based on certain semantic associations and syntactic habits; critically, one of these habits is to assume that the subject of an agent-patient verb is also the agent of the action (often called the "NVN strategy"). Many English sentences in fact show this pattern (Bever 1970). The good-enough or "satisficing" approach (e.g., Ferreira et al. 2009) similarly suggests that people often rely on heuristic-based interpretations that align with frequent patterns, such as mapping subjects to agents, during real-time comprehension. A related, neurocognitive perspective, the extended Argument Dependency Model, argues that hierarchies like animacy and thematic prominence guide alignment to shape heuristic strategies during processing (Bornkessel and Schlesewsky 2006; Bornkessel-Schlesewsky and Schlesewsky 2009). Overall, the generalizability of these heuristics isn't always clear, as they might compete with each other and with other pressures to yield language use patterns. But in general, if such experience-based strategies exist for agent-patient verbs, there may be similar strategies for other kinds of verbs, including the psychological verbs discussed in the next section.

1.2 Alignment in psychological verbs

Misalignment can be observed in some psychological verbs even when they are used in the *active* form. Unlike agent-patient verbs, which describe who does what to whom, psychological verbs describe what someone is feeling or experiencing and therefore take an experiencer and stimulus as their thematic roles (rather than an agent and patient). These are verbs like *fear, annoy, love*, and *admire* and have been studied extensively (Belletti and Rizzi 1988; Do and Kaiser 2022; Ferreira 1994; Ferstl et al. 2011; Landau 2010; Nguyen and Pearl 2021). Specifically, verbs like *fear* are *subject-experiencer* verbs, in that the subject is mapped to the experiencer in an active sentence. Conversely, verbs like

frighten are *object-experiencer* verbs, in that the *object* is mapped to the experiencer in an active sentence. Since experiencer is a more prominent thematic role than stimulus, subject-experiencer verbs are aligned, because they map the most prominent available thematic role to the subject. On the other hand, object-experiencer verbs are misaligned, because they map the less prominent stimulus to the subject despite the role of experiencer also being available.

The parallel in mapping differences between active versus passive voice and subject-experiencer versus object-experiencer verbs suggests that speakers should disprefer object-experiencers (in the active voice), just as they disprefer passives. Indeed, some converging evidence for a dispreference for object-experiencers comes from Van Gelderen (2018), who found that object-experiencer verbs were acquired one-to-two years later than subject-experiencer verbs for three English-speaking children. In general adult language use, we might expect object-experiencer verbs to be used less frequently than subject-experiencer verbs, just as passives are less frequent than actives. To assess, we would ideally compare the frequency of use of verbs like *fear* and *frighten*, which express highly similar meanings but differ in mapping. But most other psychological verbs have no counterpart that is similar in meaning but opposite in mapping between thematic role and subject. A better approach instead asks *how* we use these groups of verbs: given the mapping of a verb, are we more likely to produce it in a sentence form that creates an experiencer-as-subject? Specifically, do we show a relative preference to use subject-experiencer verbs in active sentences and object-experiencer verbs in passive sentences?

Previous work in this area has focused on comprehension, investigating how alignment between thematic and syntactic roles affects language comprehension (Ferreira 2003; Gennari and MacDonald 2008, 2009: Studies 5 and 6) and language acquisition (Nguyen and Pearl 2021). An as of yet unpublished study by Wilson and Dillon (2022) showed that mapping the experiencer to the subject facilitates understanding and results in higher grammaticality ratings compared to mapping the experiencer to the object.

A few studies suggest it is not only in comprehension but also in language production that having the experiencer as the subject matters, and that speakers will use otherwise dispreferred constructions to achieve this. In one study, when participants were asked to form sentences using two nouns and a psychological verb, they used passive voice more frequently for object-experiencer than for agent-patient and subject-experiencer verbs (Ferreira 1994). Gennari and MacDonald (2009) found a similar pattern for the construction of relative clauses. Building on Ferreira's work, Do and Kaiser (2022) showed longer speech-onset times when participants were tasked with constructing active sentences with object-experiencer verbs as compared to subject-experiencer verbs, but this pattern was reversed when the sentences had to be constructed in the passive voice.

However, in these language production studies, participants constructed sentences based on given sets of words – a task that is highly constraining in terms of free expression. When producers are free to choose the topic as well as the words and the sentence structure, the processing cost associated with object-experiencers found by previous studies may be eliminated or mitigated.

Only two prior studies have included investigation of whether subjects-as-experiencers are preferred in naturalistic contexts. Using corpora of written English, Gennari and MacDonald (2009): studies 3 and 4 computed the proportion of passives for sentences with psychological verbs (such as *annoy*) that place the experiencer in the object position in the active voice. They indeed found that these psychological verbs were used more frequently in the passive than agent-patient verbs in both main and relative clauses. In the second study, which is also the only one we know of that considered spoken language (as part of their comprehension study), Wilson and Dillon (2022) assessed the rates of passive voice use for 32 object-experiencer and subject-experiencer verbs, for a total of ~4,000 sentences, in the corpus SUBTLEX (Brysbaert and New 2009). In line with the above predictions, they found a higher rate of passive sentences for object-experiencer verbs (16.7 %) compared to subject-experiencer verbs (3.6 %). However, the SUBTLEX data represents spoken but not necessarily spontaneous language, as it comes from U.S. film and television subtitles. That at least some of the language data is scripted, the relatively small scale of the analyzed corpus, and their method of considering a wide variety of verb uses, leaves open the possibility that a spontaneous speech context might produce a pattern that is different from that observed in elicited language contexts. The present study presents a more conservative test – that is, analyzing only a restricted set of sentence types – in a large corpus of spontaneous speech.

2 Methods

To investigate whether English language relies on active-passive voice alternations to achieve experiencer-subject mappings, we conducted a corpus study of psychological verb use. We tested the hypothesis that an experiencer-as-subject pressure results in a higher passive sentence rate for object-experiencer verbs as compared with subject-experiencer verbs in spontaneous English speech (Equation (2)), regardless of the many other factors affecting verb frequency, verb type frequency, and passive sentence frequencies.

$$\frac{\text{PASSIVE OBJECT-EXPERIENCERS}}{\text{PASSIVE + ACTIVE OBJECT-EXPERIENCERS}} > \frac{\text{PASSIVE SUBJECT-EXPERIENCERS}}{\text{PASSIVE + ACTIVE SUBJECT-EXPERIENCERS}} \quad (2)$$

We selected 19 subject-experiencer and 19 object-experiencer verbs, for a total of 38 psychological verbs, which have been studied extensively in the psycholinguistics literature (e.g., Ferstl et al. 2011; Hartshorne and Snedeker 2013). The subject-experiencer verbs were *admire, adore, believe, cherish, dislike, doubt, dread, envy, favor, fear, hate, like, love, notice, pity, resent, respect, trust,* and *worship*. The object-experiencer verbs were *amaze, amuse, anger, annoy, bother, charm, disappoint, dumbfound, embarrass, fascinate, flatter, frustrate, interest, intimidate, pain, satisfy, shock, upset,* and *worry*.

We looked for uses of these verbs which were produced with two nominal arguments in either an active or passive *by*-phrase construction (e.g., like our first examples (1)-(4)). In other words, we only considered uses of these verbs which were produced with both one experiencer and one stimulus, and where both arguments were noun phrases (e.g., *The cat fascinated Ky*, *They were fascinated by all the appetizers*), which can be analyzed as a choice to map the experiencer or the stimulus to the subject. Consequently, we avoided syntactic constructions without such clear alternatives (e.g., single-argument sentences such as *They were fascinated*, or sentences with a *that*-phrase argument, such as *He believed that the cats had been out of the bag for years already*). For a single-argument sentence, it isn't obvious what the alternative is (e.g., whether *They were fascinated* should be analyzed like *They were fascinated by her*, or whether it is in fact an altogether different construction). For a *that*-phrase argument, it's doubtful that the potential alternative would ever be used (e.g., *That the cats had been out of the bag for years already was believed by him*). Our analysis thus constitutes a relatively conservative test of speakers' spontaneous choices.

We extracted all verb uses of interest from the speech section of the Corpus of Contemporary American English (COCA; Davies 2015), which are transcripts from American radio and TV programs from 1990 to 2012 (~9 million clauses). Transcripts of radio and TV conversations (typically interviews) serve as a good source of naturalistic data. COCA also contains data from fiction books, magazines, newspapers, academic texts, subtitles, blogs, and web pages, which we excluded from analysis in order to focus on speech; we also note that the current available version of the corpus also has data from more recent years than 2012. We first created a development set from one randomly-selected year of the speech section of COCA. The first and last authors manually searched through the year of transcripts for any use of the lemma strings involving any of the 38 target verbs, and compiled those uses which met the criteria for the type of verb uses of interest. In total, we created a development set containing 3,372 sentences.

We used the development set to create an automatic search for target productions from the full corpus. We tested the automatic search code by maximizing its joint recall and precision on the development set and the year of transcripts that the

development set came from. Specifically, using the open source Natural Language Processing library spaCy (Honnibal et al. 2020), we built a Python program (see supplementary materials) that returned 93.2 % of the productions in the development set and returned only 5.2 % of productions that weren't in the development set from the same year of transcripts. (That is, the program achieved 93.2 % recall and 94.8 % precision). We then ran the code on the remaining COCA transcripts to extract a total of 65,395 productions of our 38 verbs.

3 Results

The 19 subject-experiencer verbs are overwhelmingly more commonly used than the 19 object-experiencer verbs (59,730 total subject-experiencer uses, 91.3 % of the corpus vs. 5,665 object-experiencer uses, 8.7 % of the corpus). Similarly, active voice is more common than passive (64,010 active voice uses, 97.9 % of the corpus vs. 1,385 passive voice uses, 2.1 % of the corpus). Further, there are large differences in individual verb frequencies. Target uses (i.e. occurring in active or by-phrase passive sentences) of the verbs *love* and *like* each account for about 30 % of the data, with the verb *believe* accounting for about 12 %, and all remaining verbs accounting for less than 5 %. The verbs *amuse*, *pain*, *flatter*, *pity*, and *dumbfound* are very rare, each at less than 0.1 % of the total data.

As predicted, object-experiencer verbs are used in the passive (rather than active voice) much more often (19.0 % of the time) than subject-experiencer verbs (0.5 % of the time; see Table 1). However, this passivization rate varies across individual verbs (see Figure 2 and Table 2). While all subject-experiencer verbs rarely passivize, some object-experiencer verbs passivize more than half the time (e.g., *fascinate*, *intimidate*) and other object-experiencer verbs (e.g., *interest*, *worry*) passivize at lower rates comparable to the most frequently passivizing subject-experiencer verbs.

Table 1: Frequencies of active/passive voice in Subject-experiencer and Object-experiencer verb types.

Type	Voice	Count	Proportion voice/type
Subject-experiencer	Active	59,421	0.995
Subject-experiencer	Passive	309	0.005
Object-experiencer	Active	4,589	0.810
Object-experiencer	Passive	1,076	0.190

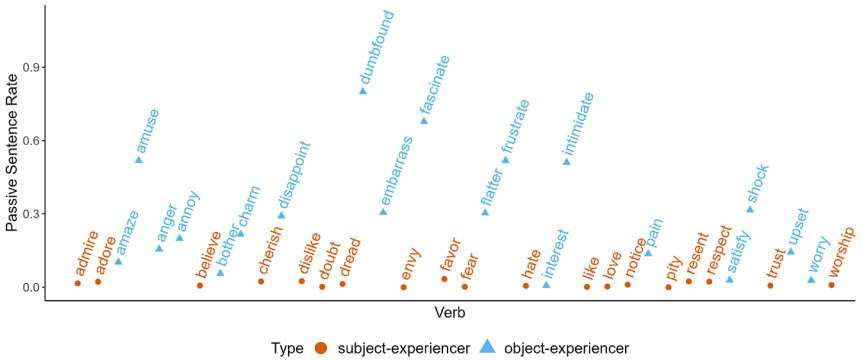

Figure 2: For each verb, the number of times it occurs in the passive out of all its uses in both the passive and active voice. Object-experiencer verbs (in blue) are more often used in the passive as compared with subject-experiencer verbs (in orange).

To test whether, despite individual variation, object-experiencer verbs are more likely to be used in the passive voice, we ran a generalized linear mixed effects model with verb type (object/subject-experiencer) predicting voice (active/passive), with random intercepts for the verb and year of its use. (The model can be found in the supplementary materials.) This model expresses the variability between verbs and takes into account the marginal probabilities of verbs within and between verb type classes.

Model results are shown in Table 3. Indeed, verb type was a highly significant predictor of voice, with object-experiencer verbs more often used in the passive voice. In order to make sense of the fixed effect coefficients, which are on the log-odds scale, we transform them to the probability scale with the inverse logit function (e.g., see Ford 2021). The predicted probability that a subject-experiencer verb is used in the passive voice is 0.8 % (the inverse logit of the intercept). Likewise the predicted probability that an object-experiencer verb is passivized is 20 % (the inverse logit of the sum of intercept and verb type: –4.7870 + 3.4032 = –1.3838). We calculated the effect size as the odds ratio, which is very high at 30.1. In other words, English speakers are thirty times more likely to use the passive voice for an object-experiencer verb as compared with a subject-experiencer verb.

4 Discussion

We asked whether English speakers prefer to present the experiencer as the subject in spontaneous uses of psychological verbs, even when this means using the generally dispreferred passive voice. Indeed, they do: to return to our original

Table 2: For each verb, the number of times it occurs in both the active and passive, and the proportion of passives out of all its uses in both the passive and active voice. Table is ordered in ascending order by passive rate.

Type	Verb	Count	Passive rate
Subject-experiencer	envy	71	0.000
Subject-experiencer	pity	21	0.000
Subject-experiencer	doubt	594	0.002
Subject-experiencer	like	19,558	0.002
Subject-experiencer	fear	557	0.002
Subject-experiencer	love	20,144	0.003
Subject-experiencer	hate	2,894	0.006
Subject-experiencer	trust	2,321	0.007
Object-experiencer	interest	141	0.007
Subject-experiencer	believe	7,635	0.007
Subject-experiencer	worship	108	0.009
Subject-experiencer	notice	1,587	0.011
Subject-experiencer	dread	72	0.014
Subject-experiencer	admire	685	0.016
Subject-experiencer	adore	310	0.023
Subject-experiencer	respect	1,565	0.023
Subject-experiencer	cherish	84	0.024
Subject-experiencer	resent	374	0.024
Subject-experiencer	dislike	237	0.025
Object-experiencer	worry	469	0.028
Object-experiencer	satisfy	272	0.029
Subject-experiencer	favor	913	0.034
Object-experiencer	bother	1,860	0.056
Object-experiencer	amaze	157	0.102
Object-experiencer	pain	44	0.136
Object-experiencer	upset	392	0.143
Object-experiencer	anger	219	0.155
Object-experiencer	annoy	111	0.198
Object-experiencer	charm	74	0.216
Object-experiencer	disappoint	196	0.291
Object-experiencer	flatter	33	0.303
Object-experiencer	embarrass	324	0.306
Object-experiencer	shock	602	0.316
Object-experiencer	intimidate	243	0.510
Object-experiencer	amuse	54	0.519
Object-experiencer	frustrate	189	0.519
Object-experiencer	fascinate	280	0.679
Object-experiencer	dumbfound	5	0.800

Table 3: Results of a generalized mixed effects model with verb type (object/subject-experiencer) predicting voice (active/passive), with random intercepts for the verb itself and the year of its use.

Fixed effects (log-odds)	Estimate	Standard error	*z*-value	*p*-value
Intercept	−4.7870	0.3068	−15.606	<2e-16
Verb type object-experiencer	3.4032	0.4184	8.134	4.14e-16

example, speakers are much happier to say *Mary is frightened by Lisa* than to say *Lisa is feared by Mary*, relative to their preference for the alternative active forms of the same sentences. Our results align with previous findings from studies of English psychological verb production (Do and Kaiser 2022; Ferreira 1994; Gennari and MacDonald 2009; Wilson and Dillon 2022).

Specifically, we found that speakers in our transcript corpus are thirty times more likely to use the passive voice for object-experiencer (e.g., *frighten*) compared to subject-experiencer verbs (e.g., *fear*). Using the passive for object-experiencer verbs maps the experiencer to the subject, a state of affairs which is already achieved when subject-experiencer verbs are used in the active voice. In other words, evidence from spontaneous language production suggests that speakers can use sentence structure to create subjects-as-experiencers, in line with a broader bias for subjects-as-experiencers.

Our study is unique in relying on language use data that is both large-scale and spontaneous. We searched through transcripts of radio and TV conversations from the Corpus of Contemporary American English (usually interviews), which serve as a good source of data on speakers' naturalistic choices. Radio and TV data allow us to consider a greater variety of speakers than psycholinguistics studies are typically able to access, and COCA in particular provides mostly unscripted speech; its web page reports that at least 95 % of the spoken section is unscripted (Davies 2024). On the other hand, a difference between the COCA data and completely natural conversation is that speakers knew that they were on a recorded program and may have altered their speech accordingly (e.g., by reducing profanity) (Davies 2024). Of note, therefore, is how our results align with the only comparable corpus study in Wilson and Dillon (2022). Despite different corpora and verbs (less than 50 % of the verbs included in the present study were also investigated by Wilson and Dillon (2022)), their study found comparable passive rates for the two types of psych verbs: 3.6 % for subject-experiencer verbs and 16.7 % for object-experiencer verbs. This suggests that despite verb-specific variation, the general picture of greater passivization for object-experiencer verbs appears to hold across the lexicon.

Nevertheless, a question for future work is the large variation in passivization rates across object-experiencer verbs. We speculate that low-passivization object-

experiencer verbs are often used in sentences besides the active and *by*-phrase passive constructions that we considered here, which nevertheless move the experiencer into the subject role, such as *bothered about* or *worried about*. For example, while *Lisa bothers **Mary*** could become ***Mary** is bothered by Lisa* in order to move the experiencer into the subject role, it could also become ***Mary** is bothered about Lisa* to express the experiencer as the subject. Similarly, *Lisa worries Mary* could become *Mary is worried about Lisa*. Further, the *bother about* and *worry about* alternatives may be preferred to *bothered by* and *worried by* because they are active rather than passive.

Finally, another alternative construction, which if considered may shed light on mapping choices, is the single-argument construction (e.g., *Lisa was annoyed* rather than *Lisa was annoyed by Mary*). It is beyond the scope of the present study to undertake to separate these constructions from adjectival uses, but we speculate that many verbs would show comparable or higher rates of passivization if single-argument cases were also included. For example, the Wilson and Dillon (2022) corpus study included single-argument cases, and the comparable passivization rates which we have cited are based on excluding probable adjectival uses (subject-experiencers passivized 3.6 % of the time and object-experiencers passivized 16.7 % of the time). When including probable adjectival uses, subject-experiencers were passivized slightly more often at 5.0 % of the time, while object-experiencers were passivized much more often at 48.3 % of the time. We hope in future research to understand the use of such additional alternative constructions.

Another potential source of variation across these verbs is the extent to which the specific case was clearly a use of a psychological subject-experiencer or object-experiencer verb. For example, a verb like *shock* may describe an electric discharge rather than entailing a mental state of its object (e.g., *The frayed wire shocked me*), and a verb like *believe* may be used to assert that the subject has a propositional attitude (e.g., *He believes that all politicians lie*). We controlled for some of this variation by limiting our corpus to sentences with two nominal arguments, including no arguments headed by a *that*-phrase or a preposition. However, additional variation might be accounted for by a more fine-grained approach to verb semantics (e.g., in line with Hartshorne and Snedeker (2013), who investigated how verb semantics predicts verb implicit causality direction).

As discussed above (Section 1.1), it remains an open question why there should be mappings that we call aligned or misaligned in a language. One of the answers is that speakers use interpretation heuristics assuming frequent alignments: for example, assuming that the sentence subject is also the agent of the action (Townsend and Bever 2001). The results of our study highlight how the experiencer-as-subject preference is potentially very similar to the agent-as-subject heuristic. It's been suggested that the experiencer role is in fact agent-like, sharing certain semantic

features commonly associated with agents (Schlesinger 1992). In other words, whatever the engines behind the agent-as-subject heuristic, they may be similar to the experiencer-as-subject preference shown here. It's worth noting that some accounts of object-experiencer verbs suggest just the opposite conceptualization of agent-like arguments, though the extent of the disagreement is unclear. For example, Dowty (1991) writes that *causing* an event or change of state in another participant is one of the contributing properties for the agent 'proto-role', and in general does not include experiencing an emotion as one of the properties that contribute to the agent proto-role. On the other hand, Dowty (1991) includes *sentience* in the agent proto-role properties, and sentience is just one of the properties that we anecdotally observe in our corpus to also distinguish the experiencer (usually sentient) from the stimulus (much less often sentient).

Finally, we wish to highlight that there are other strategies besides object-experiencer passivization which can be used to express an experiencer as a subject. An interesting finding in our corpus, in line with a prediction we briefly mentioned in Section 1.2, was that subject-experiencer verbs were far more commonly used than object-experiencer verbs. This difference in use rates could reflect a lexical rather than morphosyntactic method of satisfying the experiencer-as-subject preference, without even requiring the use of the dispreferred passive form (e.g., simply use subject-experiencer verbs in active voice much more often than object-experiencer verbs in active voice). However, a study with a broader scope is needed to verify the overall rate of subject-experiencer versus object-experiencer verb uses, as the present study only considered active and by-phrase passive constructions for a preselected list of verbs. A larger study, considering a greater variety of psychological verbs and sentences, could better compare general verb type frequency. However, this comparison would need to take into account the confounding factors of the differences in the semantics and pragmatics of most subject-experiencer verbs versus most object-experiencer verbs (that is, it is rare to have two verbs like *fear* and *frighten* which are similar in meaning but display opposite syntax-semantics mapping). These meaning-related differences are likely tied up with other pressures on overall verb type frequency.

We also note that across all verb uses, the experiencer-as-subject bias likely competes with other biases in production. We speculate that ease of production is one general consideration for which bias 'wins out'. For example, when the experiencer is a longer or more difficult phrase than the stimulus, or when it references a newer/more unfamiliar referent in the discourse, speakers might nevertheless prefer to express the experiencer as the object to ease production.

These results from spontaneous language production align well with findings from typological and historical linguistics about what may also be considered lexical strategies to express the experiencer as the subject. Object-experiencer verbs may

change to subject-experiencer verbs over time, but not vice versa (van Gelderen 2014). Further, lexicalization patterns may also be shaped by the subject-as-experiencer bias: languages like Atsugewi (Talmy 2000), Israeli Sign Language (Meir et al. 2007), and Sign Language of the Netherlands (Oomen 2017) appear to have no verbs that in their base form map the experiencer to the object (although Atsugewi allows adding a causative affix to subject-experiencer verbs to turn them into object-experiencer verbs), and American Sign Language and Japanese have numerous subject-experiencer verbs but only few object-experiencer verbs (Frederiksen and Mayberry 2021; Hartshorne et al. 2010). From this perspective, English displays an unusual pattern of more unique object-experiencer than subject-experiencer verbs (Levin 1993). Crucially, however, our results suggest that this unusual lexical pattern is mitigated in everyday language use, where English speakers can in certain cases rely on morphosyntactic means in order to achieve the preferred experiencer-subjects in actual language use.

Research funding: This research was funded in part by National Science Foundation grant no. 2005246 to Anne Therese Frederiksen. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
Data availability: The Python scripts (used to create the corpus from the transcripts in the Corpus of Contemporary English) and the R script (to analyze the corpus) are openly available at the Open Science Framework: https://osf.io/z5hb3/?view_only=ebe3c498b52f4fcba6ed53bd5413dca5.

References

Baker, Mark Cleland. 1988. *Incorporation: A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Baker, Mark Cleland. 1997. Thematic roles and syntactic structure. In Liliane Haegeman (ed.), *Elements of grammar: Handbook in generative syntax*, 73–137. Dordrecht: Springer Netherlands.

Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and -theory. *Natural Language & Linguistic Theory* 6(3). 291–352. http://www.jstor.org/stable/4047649.

Bever, Thomas G. 1970. The cognitive basis for linguistic structures. In John R. Hayes (ed.), *Cognition and language development*, 277–360. New York: Wiley & Sons, Inc.

Bock, J. Kathryn. 1982. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review* 89(1). 1–47.

Bornkessel, Ina & Matthias Schlesewsky. 2006. The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review* 113(4). 787.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass* 3(1). 19–58.

Brysbaert, Marc & Boris New. 2009. Moving beyond kučera and francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for american English. *Behavior Research Methods* 41(4). 977–990.

Davies, Mark. 2015. Corpus of contemporary American English (COCA). Version V2 https://doi.org/10.7910/DVN/AMUDUW.

Davies, Mark. 2024. *Notes on the naturalness and authenticity of the language from these transcripts*. https://www.english-corpora.org/coca/help/spoken.asp.

Dick, Frederic & Jeffrey L. Elman. 2001. The frequency of major sentence types over discourse levels: A corpus analysis. *The Newsletter of the ecenter for Research in Language, University of California San Diego* 13(1). 3–18.

Do, Monica L. & Elsi Kaiser. 2022. Sentence formulation is easier when thematic and syntactic prominence align: Evidence from psych verbs. *Language, Cognition and Neuroscience* 37(5). 648–670.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.

Ferreira, Fernanda. 1994. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language* 33(6). 715–736.

Ferreira, Fernanda. 2003. The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47(2). 164–203.

Ferreira, Fernanda, P. E. Engelhardt, M. W. Jones, N. Taatgen, H. Rijn, J. Nerbonne & L. Schomaker. 2009. Good enough language processing: A satisficing approach. *Proceedings of the 31st annual conference of the cognitive science society*, 31, 413–418. Austin, TX: Cognitive Science Society.

Ferstl, Evelyn C., Alan Garnham & Christina Manouilidou. 2011. Implicit causality bias in English: A corpus of 300 verbs. *Behavior Research Methods* 43(1). 124–135.

Ford, Clay. 2021. Getting started with binomial generalized linear mixed models. https://library.virginia.edu/data/articles/getting-started-with-binomial-generalized-linear-mixed-models.

Frederiksen, Anne Therese & Rachel I. Mayberry. 2021. Implicit causality biases and thematic roles in american sign language. *Behavior Research Methods* 53(5). 2172–2190.

Gennari, Silvia P. & Maryellen C. MacDonald. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 58(2). 161–187.

Gennari, Silvia P. & Maryellen C. MacDonald. 2009. Linking production and comprehension processes: The case of relative clauses. *Cognition* 111(1). 1–23.

Gibson, Edward, Bergen Leon & Steven T. Piantadosi. 2013. Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences* 110(20). 8051–8056.

Gordon, Peter & Jill Chafetz. 1990. Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition* 36(3). 227–254.

Grohmann, Kleanthes K., Akemi Matsuya & Eva-Maria Remberger. 2021. *Passives cross-linguistically: theoretical and experimental approaches*, Vol. 17. Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004433427.

Hartshorne, Joshua K. & Jesse Snedeker. 2013. Verb argument structure predicts implicit causality: The advantages of finer-grained semantics. *Language & Cognitive Processes* 28(10). 1474–1508.

Hartshorne, Joshua K., Timothy J. O'Donnell, Yasutada Sudo, Miki Uruwashi & Jesse Snedeker. 2010. Linking meaning to language: Linguistic universals and variation. *Proceedings of the Annual Meeting of the Cognitive Science Society* 32. 1186–1191.

Honnibal, Matthew, Ines Montani, Sofie Van Landeghem & Adriane Boyd. 2020. spaCy: Industrial-strength Natural Language Processing in Python. https://doi.org/10.5281/zenodo.1212303.

Keenan, Edward L. & Matthew S. Dryer. 2007. Passive in the world's languages. In Timothy Shopen (ed.), *Language typology and syntactic description*, 2nd edn., Vol. 1, 325–361. Cambridge: Cambridge University Press.

Landau, Idan. 2010. *The locative syntax of experiencers* (Linguistic Inquiry Monographs). Cambridge, MA/ London: MIT Press.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation*. United Kingdom: University of Chicago press.

Meir, Irit, Carol A. Padden, Mark Aronoff & Wendy Sandler. 2007. Body as subject. *Journal of Linguistics* 43(3). 531–563.

Nguyen, Emma. 2021. *The predictive power of lexical semantics on the acquisition of passive voice in young children*. Mansfield, CT: University of Connecticut dissertation.

Nguyen, Emma & Lisa Pearl. 2021. The link between lexical semantic features and children's comprehension of English verbal be-passives. *Language Acquisition* 28(4). 433–450.

Oomen, Marloes. 2017. Iconicity in argument structure: psych-verbs in sign language of The Netherlands. *Sign Language and Linguistics* 20(1). 55–108.

Plunkett, Kim & Virginia Marchman. 1993. From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48(1). 21–69.

Roland, Douglas, Frederic Dick & Jeffrey L. Elman. 2007. Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language* 57(3). 348–379.

Schlesinger, Izchak M. 1992. The experiencer as an agent. *Journal of Memory and Language* 31(3). 315–332.

Talmy, Leonard. 2000. *Toward a cognitive semantics: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Townsend, David J. & Thomas G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

van Gelderen, Elly. 2014. Changes in psych-verbs: A reanalysis of little v. *Catalan Journal of Linguistics* 13. 99–122. https://raco.cat/index.php/CatalanJournal/article/view/292445.

Van Gelderen, Elly. 2018. *The diachrony of verb meaning: Aspect and argument structure*. New York/London: Routledge.

Wagers, Matthew W. & Emily Pendleton. 2016. Structuring expectation: Licensing animacy in relative clause comprehension. *Proceedings of the 33rd west coast conference on formal linguistics*, 29–46. Cascadilla: Somerville, MA.

Wells, Justine B., Morten H. Christiansen, David S. Race, Daniel J. Acheson & Maryellen C. MacDonald. 2009. Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology* 58(2). 250–271.

Wilson, Michael & Brian Dillon. 2022. Alignment between thematic roles and grammatical functions facilitates sentence processing: Evidence from experiencer verbs. *PsyArXiv Preprints*.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cllt-2024-0091).