

What can *wh*-questions tell us about real-time language production: Evidence from English and Mandarin

Monica Do, Elsi Kaiser & Pengchen Zhao*

Abstract. We present two visual-world eye-tracking experiments investigating how speakers *begin* structuring their messages for linguistic utterances, a process known as linguistic encoding. Specifically, we focus on when speakers first linearize the abstract elements of their messages (positional processing) and when they first assign a grammatical role to those elements (functional processing). Experiment 1 decoupled the process of linearization from grammatical role assignment using *English* object *wh*-questions, where the subject is no longer sentence initial. Experiment 2 used *Mandarin* declaratives and questions, which have the same word order, to test the extent to which findings from Experiment 1 were linked to information focus associated with *wh*-questions. We find evidence of both grammatical role assignment and linearization emerging around 400-600 ms, but we do not find evidence of the +/- *wh* distinction influencing eye-movements during that same time window.

Keywords. sentence production; visual world eye-tracking; *wh*-questions; linguistic encoding; sentence formulation

1. Introduction. Language production is understood to be multi-stage and incremental, meaning that although production proceeds in a step-by-step fashion, speakers do not have to wait until every step of production is completed before starting their utterances (e.g., Schriefers et al., 1998; Ferreira and Swets, 2002; Levelt, 1989). In fact, speakers tend to plan only a small chunk of their utterances before speaking; the rest is planned “on-the-fly (Levelt, 1989; Bock and Levelt, 1994; Schriefers et al., 1998).

The messages that speakers intend to communicate are made up of pre-linguistic content that must be translated into an utterance. This process of moving from abstract, unstructured content to sequentially produced strings is known as linguistic encoding. It is at this level of production that the individual elements of the message must be assigned a grammatical function (e.g. subject, object, etc.) and a position in the sentence. These processes are known as functional and positional encoding, respectively (Bock and Levelt, 1994; Garrett, 1980, 1988).

However, an open question concerns not only the *starting point* for linguistic encoding, but also the interaction between the processes of functional and positional encoding. For instance, is linguistic encoding primarily driven by the assignment of grammatical roles (functional encoding processes) or word ordering operations (positional encoding processes)? Because work in English has predominantly focused on the production of simple transitive declaratives (but see Momma et al., 2018 for work on unaccusatives), delineating between functional and positional encoding processes has been challenging precisely because both processes are predicted to begin with planning of the subject. Consequently, work in English has led to conflicting results: Some work (Gleitman et al., 2007; Brown-Schmidt and Konopka, 2008; Myachykov and Tomlin, 2008; Myachykov et al., 2011, 2012; Tomlin, 1995, 1997) suggests the first step in linguistic

* We thank Rand Wilcox, Stefan Keine, Andrew Simpson, Victor Ferreira, Sarah Brown-Schmidt, and the USC Language Processing Lab for their feedback on this work. We also thank the audiences at AMLAP 2017, CAMP 2017, and LSA 2018 where portions of this work were presented. Authors: Monica Do (monicado@usc.edu), Elsi Kaiser (emkaiser@usc.edu), Pengchen Zhao (pengchez@usc.edu); University of Southern California.

encoding is to directly encode the most salient lexical concept as the linearly-initial item in the sentence. In a language like English, where the subject typically appears sentence-initially, that lexical concept becomes the *de facto* subject of the sentence. Other work, by contrast, has suggested that rather than assignment to a positional slot in the sentence, individual concepts are directly assigned to *grammatical roles*, beginning with the role of the subject (e.g. Griffin and Bock, 2000; Lee et al. 2013; Bock et al., 2003, 2004). For instance, Griffin and Bock (2000) used a visual-world eye-tracking paradigm to show that speakers first fixate the subject of the sentence regardless of whether the sentence is active ('The dog chased the mailman.') or passive ('The mailman was chased by the dog.'). In other words, speakers linguistically encode the sentence of the subject first, regardless of whether the subject is the semantic agent (i.e. 'doer' of the action) or patient (i.e. the entity the action is done to).

At the same time, results from cross-linguistic work using flexible word order (e.g. work by Hwang and Kaiser (2014) in Korean; Myachykov et al. (2010) in Finnish, and Myachykov and Tomlin (2008) in Russian) and verb-initial languages (e.g. see Norcliffe et al. (2015) for work in Tzeltal and Sauppe et al. (2013) for Tagalog) has been similarly difficult to interpret. Results have varied between languages (see Myachykov et al., 2011 for review) and have often not directly addressed potential effects of discourse-pragmatic factors (see Sekerina, 1997; Kaiser and Trueswell, 2004 for discussion) or have been complicated by morphological considerations (see Norcliffe et al., 2015).

In order to shed further light on the process of linguistic encoding – an in particular, on how the language production system coordinates functional versus positional encoding processes – the current work uses the visual-world eye-tracking paradigm to investigate the real-time production of declaratives and object *wh*-questions ('*Which nurses did the mailmen photograph?*') in two typologically different languages. Prior work in language production has suggested that the relevant time window for linguistic encoding in a visual-world paradigm typically occurs 400-800 ms after speakers see the image they are going to describe. Consequently, we focus our analyses and discussion on the period immediately surrounding this time window (e.g. from target image onset to 1000 ms after image onset), though we do track speakers' eye-movements through the end of each utterance.¹

2. Experiment 1: Declaratives and *wh*-Questions in English. Unlike prior work in production, which has typically focused on the production of declaratives, this work investigates the real-time production of object *wh*-questions. Experiment 1 focuses on these structures in English because the object of these sentences is obligatorily preposed to the sentence-initial position (1a-b). Crucially, because the grammatical subject of object *wh*-questions is no longer the linearly initial item in the sentence, we are able to observe the process of linguistic encoding when effects of functional encoding (e.g. subjecthood assignment) are divorced from the effects of positional encoding (e.g. linear word order).

(1a) Declarative:	The mailmen	photographed	the nurses.
	Subject	Verb	Object
(1b) Object <i>wh</i> -Question:	Which nurses	did	the mailmen photograph?
	Object	Subject	Verb

2.1. PREDICTIONS. If the starting point of linguistic encoding is driven by positional encoding processes, we expect that the sequence of linguistic encoding in English declaratives versus object *wh*-questions should differ as a result of differences in their linear word orders. In particular,

¹ For these details, we refer the reader to Do and Kaiser (under revision).

if the first step in linguistic encoding is to assign a concept to the linearly-initial slot in the sentence, then we expect that speakers should first plan the subject of declaratives, but the object in object *wh*-questions. However, if the starting point of linguistic encoding is driven by grammatical function assignment – namely, subject assignment – speakers should encode the subject first regardless of whether they ultimately produce a declarative or object *wh*-question.

2.2. PARTICIPANTS. Data from 30 native speakers of American English, recruited from the University of Southern California, were submitted for analysis.

2.3. MATERIALS AND DESIGN. Participants were asked to produce either a declarative statement or object *wh*-question about images that were presented to them on a computer screen. In order to encourage participants to ask an object *wh*-question (e.g. rather than a subject *wh*-question such as ‘Which mailman photographed the nurses?’ or a yes/no question such as “Did the mailmen photograph the nurses?”), participants told to ask a question about the characters that the action was happening to; they were also given only object *wh*-questions in example and practice items. Participants knew when to produce declaratives versus questions based on the cue that was presented to them prior to each to-be-described image: An ‘S’ preceding the target image indicated that participants should produce a statement about the upcoming image while a ‘Q’ indicated that participants should produce a question.

Each target image consisted of two sets of characters, which were left/right balanced and appeared as the subject of the sentence an equal number of times. Every image also included an instrument object which denoted the verb that participants should produce in their sentence (e.g. Koenig, 2003; McRae et al., 2005; Sussman, 2006). For instance, the camera (Figure 1) signaled to participants that they should produce the verb ‘photographed’. Verbs were chosen such that participants could not predict which characters were likely to be the subject/object of the action based only on the verb. Instead, participants were told to use the *location* of the verb-denoting instrument to determine subject/objecthood: Characters closer to the verb instrument were the subjects of the sentence. A sample image is shown in Figure 1.

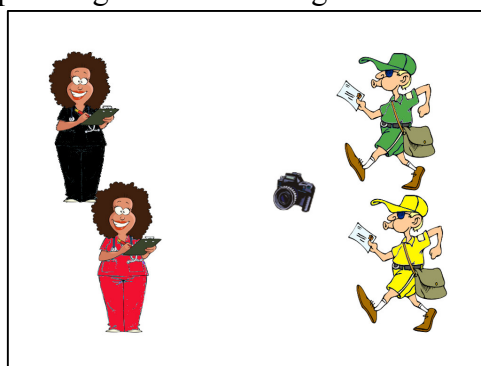


Figure 1 Sample 'mailman photographing nurses' item. Participants produced ‘*The mailmen photographed the nurses.*’ if they had seen an ‘S’ on the previous screen; they produced ‘*Which nurses did the mailmen photograph?*’ if they had seen a ‘Q’ immediately prior.

Participants were instructed to name all and only the names of the characters/objects on the screen and to start their sentences as quickly as possible. When they had finished saying their sentence out loud, they used a button on a game controller to advance to the next screen.

We also included thirty unrelated filler items for which participants were required to produce either declarative or interrogative sentences.²

2.4. PROCEDURE. Participants’ eye-movements were recorded using an Eyelink II eye-tracker (SR Research), sampling at 500Hz. Participants’ utterances were also recorded. Before moving on to the main experiment, participants learned the names for each of the characters and verb-denoting instruments used throughout the experiment.

2.5. RESULTS. Trials in which participants failed to produce the correct sentence type or used the wrong names for the images were excluded from analysis. We computed utterance onset times for each trial using Praat (Boersma and Weenink, 2009). Trials where participants took too long to begin their sentences were excluded based on the Mad-Median rule (Wilcox, 2012). These exclusions affected 15.43% of the data. To compare the strength of the subject-preference across sentence types, we calculated Subject-Object Advantage Scores. Following prior work, these Advantage Scores were calculated by subtracting the proportion of looks to the object from the proportion of looks to the subject at each time point (e.g. Kaiser, 2011; Arnold et al., 2000; Arnold et al, 2007). Those were analyzed in R (version 3.3.2; R Core Team, 2016) using Cumulative Linked Mixed Models within the ordinal package (Christensen, 2015). In all cases, effects where $|z| > 1.96$ were judged to be significant.

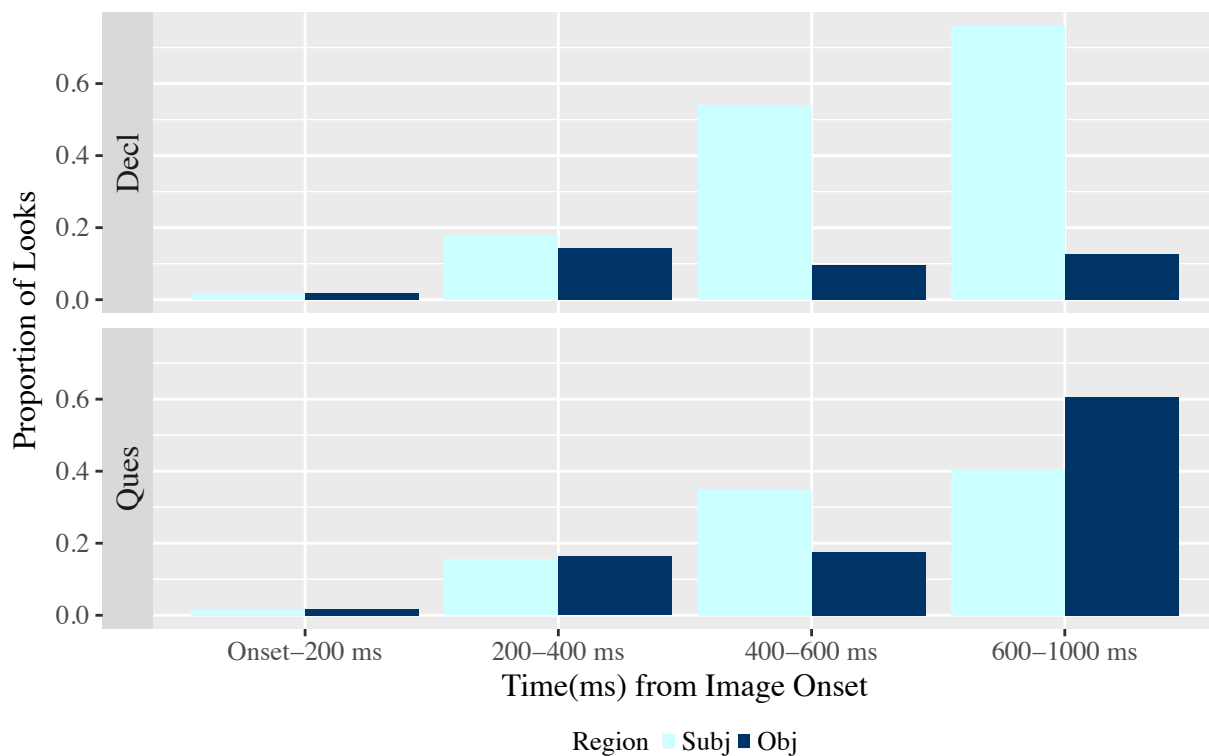


Figure 2 **English:** Proportion of looks to the Subject (light) and Object (dark) characters from Image Onset. In all time windows, $SE < .003$.

² We also included a separate interference word manipulation. Because those results are not relevant to the interpretation of the eye-movement data presented here, we do not discuss them further. For further details, see Do and Kaiser (under revision).

Note that we are primarily interested in the beginning points of linguistic encoding – in other words, *well before* speakers even begin uttering their sentences. Thus, Figure 2 plots the proportion of looks to the subject versus object characters for the first 1000 ms after critical image onset.³ People usually start to speak around 1200 ms after image onset, but this is beyond the window of interest for the starting point of linguistic encoding. Thus, what we focus on in our analyses are the moments before people start to utter the subject of the sentence.

During the first 200 ms after image onset, we find essentially no looks to either the subject or the object; this is true for both declaratives and object *wh*-questions. This is expected, because it takes approximately 200 ms to program and launch an eye-movement (e.g. Matin et al., 1993). In the 200-400 ms time window, the proportion of looks to the subject and object also show a comparable proportion of looks to the subject and object across sentence types; this is confirmed statistically, as we find no difference declaratives and questions during this time window. After the 200-400 ms time window, however, we begin to see broad differences in eye-movements for declaratives versus questions. **In declaratives**, we find that speakers consistently show a larger proportion of looks to the subject from the 400-600 ms time window onward. More specifically, looks to the subject steadily increase beginning around 400 ms, while the proportion of looks to the object remain unchanged: Across all time windows, the proportion of looks to the object are below .15. **In questions**, the proportion of looks to the subject increases in both the 400-600 and 600-1000 ms time windows, similar to declaratives. However, we find that in contrast to the eye-movements for declaratives, looks to the object *increase* in questions starting from the 400-600 ms time window and clearly overtake looks to the subject by the 600-1000 ms time window. Overall, the bar chart reflects a larger difference (i.e. a greater subject advantage) between the looks to the subject (light) versus object (dark) in declaratives than in questions. We assessed this subject advantage statistically and confirmed what is visible in the bar charts: There is a marginally significant difference between declaratives and questions at 400-600 ms that reaches significance by the 600-1000 ms window ($|z| > 2$).

2.6. DISCUSSION. Prior work in production has tended to focus on declarative, transitive sentences in English, where the subject of the sentence is also the linearly initial argument (cf. Momma et al., 2018). The current work uses English object *wh*-questions, where the object linearly precedes the subject (e.g. ‘*Which nurses did the maids tickle?*’) to investigate whether the language production system linearizes the elements an utterance before assigning grammatical roles to those elements or vice versa.

We find that in declaratives, speakers almost exclusively look to the subject immediately after image onset (we find very few looks to the object). Meanwhile, in object *wh*-questions, speakers *initially* prefer the subject (e.g. there is a subject advantage in the 400-600 ms time window), but this subject preference is short-lived: Looks to the object increase by the 600-800 ms window. Thus, one main finding of Experiment 1 is the fundamental differences speakers’ eye-movements in declaratives versus object *wh*-questions. We suggest that this difference reflects the way that speakers are able to resolve different demands arising from the linearization problem in language production. In declaratives, for instance, the linearization problem can be resolved straightforwardly: Because the subject is the first element of a declarative utterance, speakers can freely choose either to first linearize the elements of an utterance or to assign a grammatical role to those elements. In either case, the subject of the sentence will be encoded

³ We do not plot here proportion of looks to the verb because the verb is not central to our predictions. For additional information and discussion, see Do and Kaiser (under revision).

first. In line with this, we find that speakers quickly direct their gaze to the subject in declaratives, while the proportion of looks to the object remains consistently low. Object *wh*-questions, by contrast, force the language production system to weigh the need to linearize the object of the sentence first (e.g. because it appears sentence-initially) against the need to assign a syntactic subject to the message. The competition between these demands is reflected in speakers' eye-movements beginning around 400 ms when we find that the proportion of looks to the subject and looks to the object rise in parallel.

It is important to note, however, that although we did find looks to the subject and object increasing in parallel in object *wh*-questions, our results also provide evidence that grammatical role assignment (functional processing) may play a privileged role in linguistic encoding. In particular, speakers' eye-movements show both an initial preference for and a higher proportion of looks to the subject during the time window typically associated with linguistic encoding. These results suggest that linguistic encoding can start with grammatical role assignment, even when this conflicts with the linear word order of the utterance. More broadly, though, the earlier emergence of the looks to the subject suggests that the first step in linguistic encoding may be grammatical role assignment. In line with what is predicted by a functional-processing first account, speakers turn their attention to the subject even when the subject is not in the linearly-initial position.

3. Experiment 2: Declaratives and *wh*-Questions in Mandarin Chinese. Experiment 1 showed that during the time window typically associated with linguistic encoding, speakers consider the subject of the sentence before turning their attention to the object of the sentence. This is reflected by the fact that looks to the object only begin to increase only after the 400-600 ms time window. An open question not addressed by Experiment 1, though, is what drives those increased looks to the object in *wh*-questions.

One possibility is that increasing looks to the object may be driven by the word order of the utterance. In other words, speakers turn their attention to the object during linguistic encoding to assign the object a linear position in the sentence (e.g. positional processing). However, *wh*-phrases have been analyzed as discourse-pragmatically privileged, informationally focused elements. A different possibility, then, may be that the increased looks to the object in object *wh*-questions could be attributable to focus, rather than positional processing effects. Thus, the primary goal of Experiment 2 is to investigate the extent to which the results in Experiment 1 may have been influenced by information focus associated with *wh*-questions.

To tease apart the role of word order and information focus in driving looks to the object in Experiment 1, we conducted a parallel experiment in Mandarin Chinese. We choose Mandarin Chinese because of two crucial properties: First, Mandarin is a canonically Subject-Verb-Object (SVO) language, where the subject appears sentence initially. Because of this, the process of linearization in Mandarin – like declaratives in English – do not force the language production system to choose between linearization and grammatical role assignment. Second, declaratives and object *wh*-questions share the same SVO word order (2a-b). This allows us to observe potential information focus effects when constraints such as grammatical role assignment and linear word order are held constant across sentence types.

(2a) Mandarin Declarative:	救生员们	采访了	护士。
	The lifeguards	interviewed	the nurses
(2b) Mandarin <i>Wh</i> -Question:	救生员们	采访了	哪个护士?
	The lifeguards	interviewed	which nurses?

A secondary advantage of extending our work to Mandarin Chinese was to explore whether the production of object *wh*-questions in a *wh- in situ* language, such as Mandarin Chinese, would pattern like in English, where the *wh*-dependency is formed overtly. Specifically, a growing body of work (e.g. Xiang et al., 2014, 2015; Aoshima et al., 2004, Ueno and Kluender, 2009) has provided experimental evidence to corroborate prior theoretical claims (e.g., Aoun and Li, 1993; Huang, 1982; Cheng, 1991) that the same cognitive processes involved in the comprehension of *overtly* constructed dependencies also apply to covert ones. For instance, prior work by Xiang et al. (2015) using a self-paced reading paradigm found that object *wh*-phrases in Mandarin were subject to the same types of interference effects known to affect the processing of overtly constructed *wh*-dependences in languages like English (Pearlmutter et al., 1999; Van Dyke and Lewis, 2003; Lewis and Vasishth, 2005; Wagers et al., 2009; etc). In line with what work in comprehension has shown, then, we expect that the same cognitive processes to be involved in the production of overt and covert dependencies.

3.1. PREDICTIONS. One possibility is that the increased looks to the object in Experiment 1 were at least, partially driven by information focus associated with object *wh*-questions, rather than by word order – that in addition to linearization and grammatical role assignment, the process of going from thoughts to message may also be influenced by discourse-pragmatic constraints. If this is the case, we expect Mandarin object *wh*-questions to show a similar increase in looks to the object during the same time windows – roughly 600-1000 ms after image onset. If we, by contrast, do not find increased looks to the object during the 600-1000 ms windows, this would suggest that the pattern of results found in Experiment 1 are unlikely to be linked to information focus effects.

3.2. PARTICIPANTS. 35 native speakers of Mandarin Chinese, all born and raised in mainland China, were recruited from the University of Southern California.

3.3. MATERIALS AND DESIGN. Materials and design used in Experiment 2 were the same as in Experiment 1. To account for lexical differences between languages, some lexical items – such as characters, verbs, and the sentence-type indicator – were modified.

3.4. PROCEDURE. The procedure was the same as Experiment 1.

3.5. RESULTS. In Experiment 2, 22.59% of the data were excluded due to errors and disfluencies in the utterances. Using the Mad-Median rule, we also excluded trials where speakers took too long to begin their utterances; this affected 8.03% of the overall data. Data from Experiment 2 was analyzed using the same statistical methods as in Experiment 1.

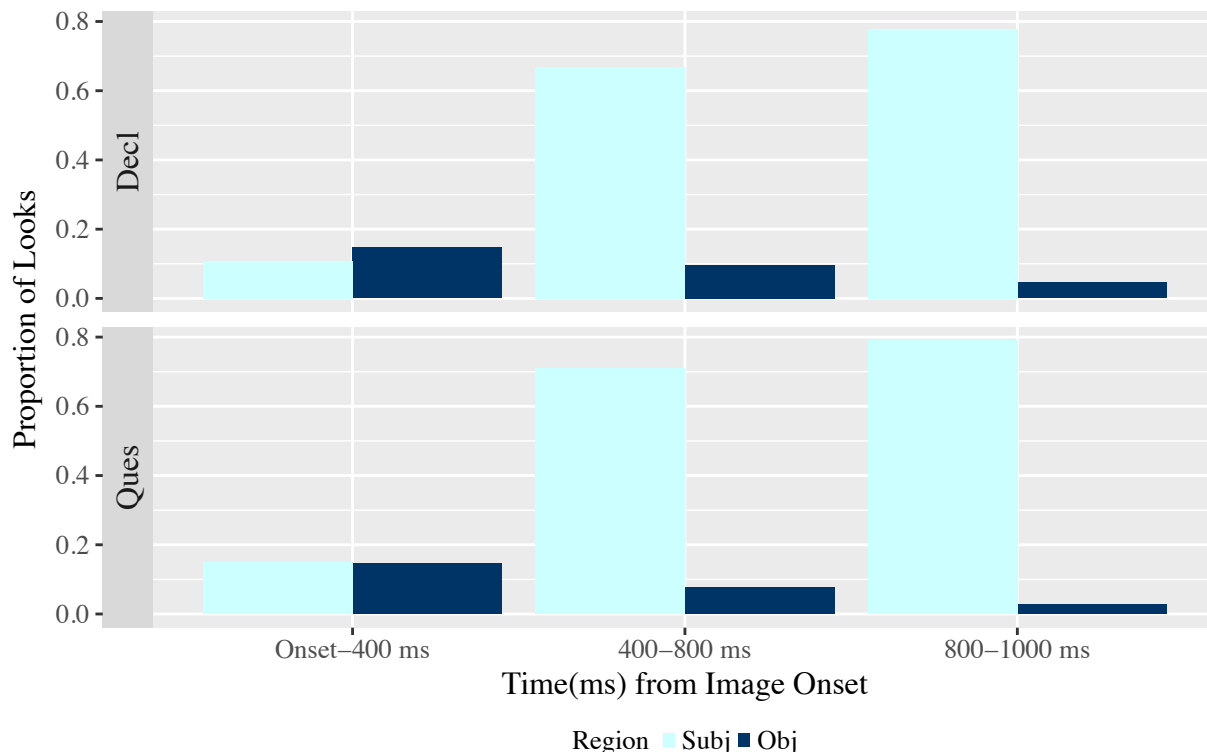


Figure 3 **Mandarin Chinese**: Proportion of looks to the Subject (light) and Object (dark) characters from Image Onset. In all time windows, $SE < .002$.

As in English, we find essentially no looks to either the subject or object in the first 200 ms after image onset. This is predicted given that it takes 200 ms to program and launch an eye-movement (Matin et al., 1993). After 200 ms, declaratives and questions largely show the same pattern: Looks to the subject and object increase during the 200-400 ms time window for both sentence types. After 400 ms, we see a sharp increase in the proportion of looks to the subject in both declaratives and questions; meanwhile, looks to the object remain relatively low. To directly assess the strength of the subject preference across sentence types in Mandarin Chinese, we again compute subject advantage scores. These scores confirm what is visually present in the graphs: We find no differences in eye-movements in declaratives and object *wh*-questions in any of the time windows after image onset.

3.6. DISCUSSION. The goal of Experiment 2 was to investigate the possibility that the eye-movements from Experiment 1 were driven, to some degree, by informational focus associated with object *wh*-questions rather than by linear word order effects. Experiment 2 showed that when the linear word order of declaratives and object *wh*-questions is the same, as in Mandarin Chinese, no differences in speakers' eye-movements during the first 1000 ms of production. These results suggest, therefore, that the pattern of results obtained in Experiment 1 are unlikely to be driven by information focus effects associated with object *wh*-questions.

4. General Discussion. Even though questions are a familiar and common part of everyday communication, this is the first work to examine the real-time production of questions. We did this in two typologically different languages, English (Experiment 1) and Mandarin Chinese (Experiment 2). The main aim of our work in English (Experiment 1) was to get a better view of the process of linguistic encoding during sentence production. In particular, prior work has sug-

gested that going from abstract concepts to linguistic representations requires the language production system to assign each concept with (i) a grammatical function, a process known as functional encoding and (ii) a positional slot in the utterance, a process known as positional encoding.

The goal of this work was to shed further light on how these processes are coordinated within the level of linguistic encoding. Taken together, our results contribute to this question in two ways. First, functional and positional encoding appear to occur in tandem, suggesting that both factors play a role during the process of linguistic encoding. This conclusion is informed by the fact that in object *wh*-questions (when the subject is no longer the linearly initial word in the sentence), speakers look at both the subject and object characters during the same general time window. This is in contrast to declaratives, where speakers attend strictly to the subject. Nonetheless, we also find evidence that functional processes play a privileged role during linguistic encoding: We find a significantly greater proportion of looks to the subject than to the object. Based on the results from Experiment 2 (Mandarin), we confirmed that the results in English were unlikely to be due to information focus effects between declaratives and questions and instead stem from the syntactic differences between them.

Second, our work suggests the way in which speakers move from abstract ideas to linguistic utterances may be more sensitive to syntactic structure than one might expect. For instance, even though the end product of sentence production superficially appears to be a linear string, an ‘intuitive’ approach to linguistic encoding could be to follow the linear order in which each concept needs to be uttered. The results presented here, though, argues for a model of linguistic encoding that is not only *able* to engage in non-linear planning, but is *immediately* sensitive to the underlying syntactic structure of the utterance at hand.

More broadly, the picture that has begun to emerge is one in which the process of production is influenced by a number of factors. At the level of message formulation, for instance, other work has shown that multiple factors – including perceptual salience, conceptual salience, and the complexity of the message – may compete to influence the starting point of message formulation. This is also true of work in the domain of lexical access, which has shown interactive effects of word frequency, semantic similarity, etc. The results of the current work also provide evidence for a flexible and multi-factorial approach to linguistic encoding.

References

- Aoshima, S., Phillips, C., and A. Weinberg. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51. 23–54. <https://doi.org/10.1016/j.jml.2004.03.001>.
- Aoun, J. and Li, A. Y-H. (1993). *Wh*-elements in situ: Syntax or LF? *Linguistic Inquiry* 24. 199–238.
- Arnold, J., Brown-Schmidt, S., and Trueswell, J. C. (2007). Children’s use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes* 22(4). 527–565. <https://doi.org/10.1080/01690960600845950>.
- Arnold, J., Wasow, T., Losongco, A., and Ginstrom, R. (2000). Heaviness vs. newness: the effects of complexity and information structure on constituent ordering. *Language* 76, 28–55. <https://doi.org/10.1353/lan.2000.0045>.
- Bock, J. K., Irwin, D. E., and Davidson, D. J. (2004). “Putting first things first,” in *The integration of language, vision, and action: Eye movements and the visual world*, eds J. Henderson and F. Ferreira (New York: Psychology Press), 249–278.

- Bock, K., Irwin, D. E., Davidson, D. J., and Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language* 48. 653–685. [https://doi.org/10.1016/S0749-596X\(03\)00007-X](https://doi.org/10.1016/S0749-596X(03)00007-X).
- Boersma, P., and Weenink, D. (1992–2011). Praat: Doing phonetics by computer. Retrieved from <http://www.praat.org/>.
- Brown-Schmidt, S., and Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, 109(2), 274-280. <https://doi.org/10.1016/j.cognition.2008.07.011>.
- Cheng, L. L-S. (1991). On the typology of *wh*-questions. Doctoral dissertation, MIT.
- Christensen, R.H.B. (2015). ordinal - Regression Models for Ordinal Data. R package version 2015.6-28. <http://www.cran.r-project.org/package=ordinal/>.
- Do, M.L., and Kaiser, E. (under revision). Linguistic Encoding is multi-factorial and competitive: Evidence from real-time production of *wh*-questions in two languages
- Ferreira, F., and Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language* 46. 57-84. <https://doi.org/10.1006/jmla.2001.2797>.
- Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production* (Vol. 1, pp. 177-220). London: Academic Press.
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey: III. Language: Psychological and biological aspects* (pp. 69-96). Cambridge: Cambridge University Press.
- Gleitman, L., January, D., Nappa, R., and Trueswell, J. (2007). On the give- and-take between event apprehension and utterance formulation. *Journal of Memory and Language* 57. 544–569. <https://doi.org/10.1016/j.jml.2007.01.007>.
- Griffin, Z. M., and Bock, K. (2000). What the eyes say about speaking. *Psychological Science* 11. 274-279. <https://doi.org/10.1111/1467-9280.00255>.
- Huang, C.-T. J. (1982). Logical Relations in Chinese and the Theory of Grammar, PhD dissertation, MIT.
- Hwang, H., and Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology* 40(5). 1363-1376. <https://doi.org/10.1037/a0036797>.
- Kaiser, E. (2011). Focusing on pronouns: Consequences of subjecthood, pronominalisation, and contrastive focus. *Language and Cognitive Processes* 26. 1625-1666. <https://doi.org/10.1080/01690965.2010.523082>.
- Kaiser, E. and Trueswell, J. C. (2004) The role of discourse context in the processing of a flexible word-order language. *Cognition* 94. 113-147. <https://doi.org/10.1016/j.cognition.2004.01.002>.
- Koenig, J.P., Mauner, G., and Bienvenue, B. (2003). Arguments for adjuncts. *Cognition* 89(2). 67-103. [https://doi.org/10.1016/S0010-0277\(03\)00082-9](https://doi.org/10.1016/S0010-0277(03)00082-9).
- Lee, E., Brown-Schmidt, S., and Watson, D. G. (2013). Ways of looking ahead: Hierarchical planning in language production. *Cognition* 129. 544–562. <https://doi.org/10.1016/j.cognition.2013.08.007>.
- Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lewis, R. L. and Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 375-419. https://doi.org/10.1207/s15516709cog0000_25.

- Matin E., Shao K.C., Boff K.R. (1993). Saccadic overhead: information processing time with and without saccades. *Perception and Psychophysics* 53(4). 372-380. <https://doi.org/10.3758/BF03206780>.
- McRae, K., Hare, M., Elman, J., and Ferretti, T. (2005) A basis for generating expectancies for verbs from nouns. *Memory and Cognition* 33(7). 1174-1184. <https://doi.org/10.3758/BF03193221>.
- Momma, S., Slevc, L.R., & Phillips, C. (2018). Unaccusativity in sentence production. *Linguistic Inquiry*, 49, 181–194. https://doi.org/10.1162/LING_a_00271.
- Myachykov, A., Garrod, S., and Scheepers, C. (2010). Perceptual priming of structural choice during English and Finnish sentence production. In *Language and cognition: state of the art*, ed. by R. K. Mishra and N. Srinivasan, 54–72. Munich: Lincom Europa.
- Myachykov, A., and Tomlin, R. S. (2008). Perceptual priming and structural choice in Russian sentence production. *Journal of Cognitive Science* 6(1). 31–48. <https://doi.org/10.17791/jcs.2008.9.1.31>.
- Myachykov, A., Thompson, D., Scheepers, C., and Garrod, S. (2011). Visual attention and structural choice in sentence production across languages. *Language and Linguistic Compass* 5. 95–107. <https://doi.org/10.1111/j.1749-818X.2010.00265.x>.
- Myachykov, A., Thompson, D., Scheepers, C., and Garrod, S. (2012). Determinants of structural choice in visually situated sentence production. *Acta Psychologica* 141, 304-315. <https://doi.org/10.1016/j.actpsy.2012.09.006>.
- Norcliffe, E., Konopka, A. E., Brown, P., and Levinson, S. (2015). Word order affects the timecourse of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*. <https://doi.org/10.1080/23273798.2015.1006238>.
- Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language* 41. 427-456. <http://dx.doi.org/10.1006/jmla.1999.2653>.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin, R. D., and Levinson, S. (2013). Dependencies first: Eye-tracking evidence from sentence production in Tagalog. Paper presented at the 35th meeting of the Cognitive Science Society, Berlin, Germany, July 2013.
- Schriefers, H., Teruel, E., and Meinshausen, R. M. (1998). Producing simple sentences: Results from picture–word interference experiments. *Journal of Memory and Language*, 39, 609–632. <https://doi.org/10.1006/jmla.1998.2578>.
- Sekerina, I.A. (1997). The syntax and processing of scrambling constructions in Russian. Doctoral dissertation, City University of New York, New York, NY.
- Sussman, R. (2006). Verb-Instrument Information During Online-Processing. Doctoral Dissertation. University of Rochester, Rochester.
- Tomlin, R. S. (1995). Focal attention, voice, and word order. In *Word Order in Discourse*, eds P. Downing and M. Noonan (Amsterdam: John Benjamins), 517–552.
- Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: the role of attention in grammar. In *Language and Conceptualization*, eds J. Nuyts and E. Pederson (Cambridge: Cambridge University Press), 162–189.
- Ueno, M. and Kluender R. (2009). On the processing of Japanese *wh*-questions: An ERP study. *Brain Research* 1290. 63-90. <https://doi.org/10.1016/j.brainres.2009.05.084>.

- Van Dyke, J. A., and Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49. 3: 285-316. [https://doi.org/10.1016/S0749-596X\(03\)00081-0](https://doi.org/10.1016/S0749-596X(03)00081-0).
- Wagers, M., Lau, E. and Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language* 61. 206-237. <https://doi.org/10.1016/j.jml.2009.04.002>.
- Wilcox, R. R., and Keselman, H. J. (2012). Modern regression methods that can substantially increase power and provide a more accurate understanding of associations. *European Journal of Personality* 26(3), 165-174. <https://doi.org/10.1002/per.860>.
- Xiang, M., Dillon, B., Wagers, M., Liu, F. Q., and Guo, T. M. (2014) Processing covert dependencies: an SAT study on Mandarin *wh*-in-situ questions. *Journal of East Asian Linguistics* 23(2). 207-232. <https://doi.org/10.1007/s10831-013-9115-1>.
- Xiang, M., Wang, S-P., and Cui, Y-L. (2015) Constructing covert dependencies—The case of Mandarin *wh*-in-situ dependency. *Journal of Memory and Language* 84. 139-166. <http://dx.doi.org/10.1016/j.jml.2015.05.006>.