Production predicts comprehension: Animacy effects in Mandarin relative clause processing

CrossMark

Yaling Hsiao, Maryellen C. MacDonald *

Department of Psychology, University of Wisconsin-Madison, WI, USA

ARTICLE INFO

ABSTRACT

This article investigates the relationship between production and comprehension of relative clauses in Mandarin Chinese. In a picture description task, we find strong head noun animacy effects on relative clause production despite the fact that Mandarin has head-final relative clauses ([[relative clause] head noun]), so that the animate/inanimate head noun is uttered late. These and other production results have implications for theories of incremental language planning. We then used corpus analyses to investigate the distribution of structure–message pairings in the language that result from these animacy-based production biases. Mandarin is particularly interesting from the language comprehension side, as there is an extensive literature on relative clause comprehension, with conflicting results. A gated sentence completion task reveals comprehenders' animacy-linked expectations in relative clause interpretation and also shows the substantial amount of syntactic ambiguity in Mandarin relative clauses, owing to their head-final structure. The completion data were reliable predictors of comprehenders' self-paced reading times, but the distance between syntactically-dependent elements in the sentences was not. We argue that these results argue against accounts of sentence comprehension that posit that some sentence types are inherently more difficult than others. Instead we suggest that sentence types with which comprehenders have little experience are difficult, and we link the results ultimately to producers' different production choices in different animacy configurations and consequent variation in language patterns that comprehenders experience.

© 2015 Elsevier Inc. All rights reserved.

Introduction

In order to test theories of how people comprehend language, researchers often measure comprehension difficulty of certain sentence types via some combination of reading times and accuracy, typically via responses to comprehension questions. Because differences in time/accuracy across sentence conditions cannot by themselves indicate why one sentence type is harder than another, data patterns are given different interpretations in various theoretical accounts of comprehension processes. Recent sentence processing research has been divided between experience-based approaches, including both constraint-based accounts (MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus & Trueswell, 1995) and surprisal (Hale, 2001; Levy, 2008), and inherent-difficulty accounts (Gibson, 1998, 2000; King & Just, 1991; Lewis & Vasishth, 2005; Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Johns, 2012). Briefly, experience-based approaches suggest that comprehenders gain skill at some or all levels of language interpretation from past sentence comprehension experience, so that comprehension difficulty is predicted to vary with past experience with similar language

* Corresponding author at: Department of Psychology, University of Wisconsin-Madison, Language and Cognitive Neuroscience Lab, 1202 W. Johnson St., Madison, WI 53705, USA.

E-mail address: mcmacdonald@wisc.edu (M.C. MacDonald).

input—similar sentence structures, words, word sequences, and/or structure–word pairings. By contrast, inherent-difficulty accounts suggest that some sentences are inherently harder than others because they are computationally more difficult, by virtue of having greater syntactic complexity, longer distance dependencies between related words, or other features that tax comprehenders' comprehension abilities or working memory.

Relative clauses have played an outsized role in generating evidence for both of these theoretical positions. For example, the reduced passive relative clauses in (1) are frequently given as examples of the influence of comprehenders' use of past experiences with verb–noun thematic role combinations and/or experience with events in the world (e.g., Trueswell, Tanenhaus, & Garnsey, 1994): Sentence (1a) is more difficult than (1b) because in comprehenders' past experience, animate nouns such as *girl* are plausible and common agents of verbs such as *kissed*, leading to misanalysis of the sentence, while *trophy* in (1b) is not a good agent of *kiss*, allowing the comprehender to rapidly settle on the correct analysis, in which the trophy is the object of *kissed*. On this view, past experience affects comprehenders' abilities to resolve ambiguities and settle on the correct interpretation of the input.

(1a)   The girl kissed by the woman was . . .
(1b)   The trophy kissed by the woman was . . .

Inherent-difficulty approaches have also pointed to relative clauses, specifically the contrast between Subject Relative Clauses as in sentence (2a) and Object Relative Clauses as in (2b). A large number of studies have found that object relatives (2b, in which the head of the relative clause *girl* is the object of the relative clause verb *kissed*) are harder than subject relatives (2a, where the head *girl* is the subject of *kissed*) in English and many other languages (see Gennari & MacDonald, 2008; O'Grady, 2011, for reviews). Because these sentences have been assumed to be unambiguous, the difficulty contrast has generally been thought to owe to syntactic complexity differences between the sentences. Gibson (1998), for example, argued that object relatives create a higher working memory load than do subject relatives, because the object relatives involve longer distance relationships between dependent elements, such as between a verb and its subject or object (see also Gordon, Hendrick, & Johnson, 2004; Grodner & Gibson, 2005; Hawkins, 1994, 2004; King & Just, 1991; Lewis et al., 2006; for related memory-based approaches).

(2a)   Subject Relative Clause: The girl [who kissed the woman] was . . .
(2b)   Object Relative Clause: The girl [who the woman kissed] was . . .

In the broadest sense, both inherent properties of the linguistic input and prior experience will play a role in comprehension difficulty. For example, if experimenters presented sentences at pitches higher than a person's hearing range, those materials would be inherently difficult to comprehend, and additional experience wouldn't change that fact. Similarly, researchers who claim the inherent

difficulty of certain sentence types do not suggest that prior experience is irrelevant; minimally language users must learn the words of their language, via experience. However, the existence of both experience and inherent difficulty components of language comprehension in general does not mean that a particular result, such as contrasts between two relative clause types, must necessarily have both experience and inherent difficulty components (cf. Fedorenko, Woodbury, & Gibson, 2013; Staub, 2010). Specifically for our own approach, we have argued that memory limitations are not independent of prior experience, meaning that as comprehenders gain skill with interpreting a given sentence structure (via prior experience with it, including generalization to new tokens, Hsiao & MacDonald, 2013; MacDonald & Christiansen, 2002), the computational capacity (memory) needed to interpret a new token of this sentence type declines. If this approach is correct, then we cannot calculate memory load directly from inherent properties of sentences such as the distance between words in a sentence, as in Dependency Locality Theory (Gibson, 1998) or cue-based parsing approaches (Lewis et al., 2006), because load depends on an individual comprehender's prior experience and need not be a direct reflection of distance or other sentence properties. On this view, there is no experience-independent inherent memory load—the memory-linked difficulty of a sentence is always dependent on comprehenders' prior experience, to the point that in principle, longer dependency distances could be easier than shorter ones, given relevant experience. Support for the malleability of difficulty comes from Gennari and MacDonald's (2008) finding that the distance between dependent elements in a sentence (the difficulty metric in Gibson's Dependency Locality Theory) accounted for no additional variance beyond experience-based measures. Relatedly, a number of studies manipulating relative clause experience in human comprehenders and computational models show how difficulty changes with experience with other structures in the language (Hsiao & MacDonald, 2013; MacDonald & Christiansen, 2002; Wells, Christiansen, Race, Acheson, & MacDonald, 2009).

These results argue for a strong link between experience and comprehension of relative clauses, but the situation is quite complicated and open to alternative interpretations. Extension to other languages can provide an important test of the experience and inherent difficulty accounts, not only because of the value of replication, but also because the structure of relative clauses differs widely across languages, and because relative clause comprehension has been intensely studied in a number of different languages. After English, the language for which relative clause processing has received perhaps the most scrutiny is Mandarin (e.g., Hsiao & Gibson, 2003; Hsiao & MacDonald, 2013; Jäger, Chen, Li, Lin, & Vasishth, 2015; Lin & Bever, 2006; Vasishth, Chen, Li, & Guo, 2013; Wu, Kaiser, & Andersen, 2012). As we describe below, Mandarin is intensively studied because its combination of relative and main clause structure is typologically unique, its patterns of comprehension difficulty are markedly different from those found in English and most other languages, and it has been used as a central test case for inherent difficulty, specifically a role for dependency

distance in relative clause processing (Hsiao & Gibson, 2003). Our goal in this paper is to investigate the production–comprehension linkages in Mandarin relative clauses, from speakers' choices during production, the resulting distributions in the language, comprehenders' expectations for relative clause resolution as they are in the midst of a sentence, and comprehension difficulty, as measured by reading times. We begin with a description of Mandarin relative clauses and the comprehension patterns that have placed them at the forefront of accounts of relative clause processing.

Production–comprehension links in Mandarin

Mandarin has a head-final relative clause structure, meaning that the head noun follows the relative clause, as in the examples in (3). It has an obligatory relativizer DE (similar to *that* in English), but as these examples show, *de* does not arrive until the end of the relative clause.

of subject-modifying relative clauses in (3) does not know whether these words begin a relative clause or, (much more common) a simple main clause (Hsiao & MacDonald, 2013; Jäger et al., 2015). For example, the *attack opponent* . . . sequence at the start of the subject relative in (3a) could instead be the common null-subject pro-drop structure in Mandarin, meaning that someone attacked the opponent. Similarly, the object relative in (3b) begins *opponent attack* . . ., which could be the beginning of a simple sentence in which the opponent is attacking something. Hsiao and MacDonald (2013) showed that object-modifying relative clauses afforded less ambiguity, affecting patterns of comprehension difficulty. Together, these results suggest that the effects are complex and lexically-dependent, a situation that could plausibly be clarified by studying language production and the frequency of alternative forms that are produced, and thus are in comprehenders' linguistic environment.

(3a)	Subject Relative:	[攻击 竞争者 的]	候选人	赢了	这场	选举。
		[e_1 attack opponent *de*]	candidate$_1$	won	this	election
		The candidate who attacked the opponent won this election.				
(3b)	Object Relative:	[竞争者 攻击 的]	候选人	赢了	这场	选举。
		[opponent attack e_1 *de*]	candidate$_1$	won	this	election
		The candidate who the opponent attacked won this election.				

There have been relatively few studies of relative clause production in Mandarin, but there has been extensive comprehension research, most of which has examined the comprehension difficulty of subject vs. object relatives, as in (3a–b). Some studies have found that Mandarin subject relatives are more difficult than object relatives, the reverse of patterns in most other languages (Chen, Ning, Bi, & Dunlap, 2008; Gibson & Wu, 2013; Hsiao & Gibson, 2003; Lin & Garnsey, 2011; Su, Lee, & Chung, 2007). This result finds a clear interpretation in inherent-difficulty/memory accounts. Hsiao and Gibson argued that subject relatives (e.g., 3a) were more difficult than object relatives (e.g., 3b) because they have higher storage and integration costs, owing to the longer distance between dependencies, as shown by the co-indexed filler and gap (*e*) in the examples. Hsiao and Gibson's claim is that the greater number of intervening words between the filler and the gap in subject relatives permits the introduction of more new discourse referents and incomplete dependencies, leading to higher memory loads in subject relatives than in object relatives. However, other studies have not replicated these results but have instead found the typical pattern of subject relatives being easier than object relatives (e.g. Jäger et al., 2015; Lin & Bever, 2006; Vasishth et al., 2013). Still other studies have suggested that the pattern of comprehension difficulty is a function of whether the relative clause modifies the main clause subject (as in 3a–b) or the main clause object, as in *The journalist respected the candidate who* . . . (Lin & Bever, 2006; Yang, Perfetti, & Liu, 2010). Hsiao and MacDonald (2013) traced these results to different amounts of ambiguity across sentence types. A consequence of the Mandarin head-final relative clause structure is that the comprehender who encounters the initial words

Hsiao and MacDonald (2013) presented corpus data relevant to this language production question. Their results show that rates of production of several relative clause types in Mandarin are intimately tied to animacy of the head noun. One obvious factor affecting producers' implicit structure choices is the relative accessibility of animate vs. inanimate heads, and many production researchers have observed that animate entities tend to be placed early in the sentence and/or in grammatically prominent positions, such as sentence subject, which yields a passive structure (Bock, 1987; Bock & Warren, 1985; Tanaka, Branigan, McLean, & Pickering, 2011). However, Gennari, Mirković, and MacDonald (2012), who studied relative clause production in English, Spanish, and Serbian, argued that accessibility of animate heads does not fully explain production choices. Specifically, patterns of production appear to be shaped by a bias that MacDonald (2013) called Reduce Interference. The presence of two semantically similar words in an utterance increases production difficulty (Smith & Wheeldon, 2004). This result may owe to retrieval interference—in planning an utterance about a girl and retrieving the lemma *girl* from long term memory cued by its conceptual features, it then becomes more difficult to map a similar set of conceptual features to a different lemma, *boy* (Oppenheim, Dell, & Schwartz, 2010). Gennari et al. observed that in cases where both participants in an event were animate (e.g. a boy splashing a girl), not only were there more passive relative clauses as in (4a), but there were more bare (agentless) passives like (4b), a result that cannot be explained solely by a bias to place accessible animate heads in subject position. Gennari et al. linked these effects to semantic similarity between the two nouns in the utterance, not simply animacy of a single noun, and

they hypothesized that another motivation for using a passive with an animate patient was to reduce this memory interference—the passive allows demotion of the agent to a syntactically minor *by* phrase or omission altogether, reducing this similarity-based interference.

(4a) Passive Relative Clause: The girl [who was kissed by the woman]...

(4b) Bare Passive Relative Clause: The girl [who was kissed]...

It is interesting to consider these Reduce Interference claims with respect to Mandarin. The strong animacy effects on relative clause production have typically been conducted in European languages with head-first relative clauses, where the animate/inanimate head noun precedes the relative clause. However, in Mandarin, the relative clause precedes the head noun, and so it is unclear whether this late-appearing head noun will influence the nature of the earlier relative clause. On the one hand, it is reasonable to assume that producers must plan the head noun before or at the same time as the relative clause that modifies it, even though the head noun will be uttered later. On this view, animacy and semantic similarity effects on relative clause utterance choices could be expected to be similar in Mandarin to those found in head-first relative clauses. On the other hand, the cumulative semantic interference studies that demonstrate semantic similarity effects involve overt production of words that interfere with some subsequent target word (Howard, Nickels, Coltheart, & Cole-Virtue, 2006; Oppenheim, Dell, & Schwartz, 2007) or have near-simultaneous presentation of interfering words (picture–word interference studies, e.g., Glaser & Düngelhoff, 1984; Meyer, 1996), so that these interference effects may not inform sentence production in Mandarin. Thus while we expect to find interference effects even with head-final relative clauses, a different pattern of results is also plausible.

Hsiao and MacDonald (2013) did see animacy–structure patterns in relative clauses in their corpus data, but they did not examine passive relatives like the English counterparts in (4a–b). The current study investigates the full chain of relationships from production, distribution to comprehension in Mandarin relative clauses. Experiment 1 adopted Gennari et al.'s (2012) picture description paradigm to investigate speakers' choices of relative clause form through manipulating head noun animacy. This study will therefore allow direct comparison to the languages with head-first relative clauses that Gennari et al. studied. Because studies of Mandarin relative clause comprehension have used written stimuli that may not share the distributional patterns of spoken language (Drieman, 1962;

Schallert, Kleiman, & Rubin, 1977), we compared our spoken production results to written corpus data to identify distributional patterns in speech vs. text, with the aim of developing written materials that would enable comparisons with those in the existing literature. To foreshadow our findings, we found differences between spoken and written production, and we used our and others' corpus and production findings to construct felicitous written stimuli for a "gated" sentence completion task (McRae, Spivey-Knowlton, & Tanenhaus, 1998) in Experiment 2 and a sentence comprehension study in Experiment 3. We thus aim to relate comprehension processes in Experiment 3 to comprehenders' expectations, as measured by sentence completion data in Experiment 2, which in turn stem from distributions in the language that owe to speakers' production choices, as in Experiment 1.

The relative clause contrast we investigate in these studies is the contrast between object relatives (5a) and the passive relative counterpart, (5b), which is a type of subject relative in which the verb has passive voice. There are several reasons for this comparison. First, the two alternatives are closely matched in meaning, and this is why the object vs. passive relative was the contrast of choice in Gennari and MacDonald's (2008, 2009) study of the production–comprehension link. In particular, they argued that this choice was preferable to the common practice of comparing (active voice) subject and object relatives as in (2a–b), as these sentences confound sentence structure and meaning (see also MacDonald, Montag, & Gennari, in press). Of course, active and passive forms are not fully identical, but compared to the typical contrast of active forms of subject vs. object relative with clearly different meanings, the object vs. passive relative comparison allows us to examine producers' choice between two similar alternatives when prompted by a picture. Second, Mandarin passive relative clauses do not differ from object relatives in word order; their only difference is a clause-initial passive marker *bei*, as seen in (5b). Third, passives (which are syntactically subject relatives) have a longer dependency distance than object relatives, as shown with the subscripts in (5a–b). This is a very useful property, because previous memory-based accounts of relative clause comprehension in both Mandarin and English have pointed to differences in dependency distance between words in the sentence in explaining patterns of processing difficulty in Mandarin object and (active voice) subject relatives (Gibson, 1998; Hsiao & Gibson, 2003). By choosing a contrast with similar meaning, identical word order, and different dependency distance, we can contrast the dependency claims with our own experience-based account, in which comprehension behavior reflects the distributional patterns generated by production choices that are strongly tied to lexical properties of the utterance—head noun animacy.

(5a)	Object Relative:	[竞争者 攻击 的]		候选人	赢了	这场	选举。
		[opponent attack e_1 *de*]		candidate$_1$	won	this	election
		The candidate who the opponent attacked won this election.					
(5b)	Passive relative:	[被 竞争者 攻击 的]		候选人	赢了	这场	选举。
		[e_1 *bei* opponent attack *de*]		candidate$_1$	won	this	election
		The candidate who was attacked by the opponent won this election.					

Experiment 1 – picture description task

This experiment investigated the production of active vs. passive relative clauses in Mandarin and the degree to which head noun animacy affected speakers' production choices. We adopted Gennari et al.'s (2012) visual scene description paradigm and used a majority of their materials. Some new experimental pictures were developed to replace ones that did not work well in Mandarin (e.g. because different verbs are used for actions on animate vs. inanimate entities). Because the active and passive relative clauses have identical word order in Mandarin (distinguished only by the clause-initial passive marker *bei* in the passive relative clause), any preference for animate passives cannot be attributed to preferences for a certain surface word order (as arises in some production environments, e.g. Tanaka et al., 2011) but rather a preference for a particular grammatical role assignment (e.g. Bock & Warren, 1985): animate entities are more salient and usually occupy the grammatical subject position. By using passives, speakers could align the animate nouns as subject in the relative clause. Following results in Spanish and English, we expect that head noun animacy/accessibility will affect rates of passive relatives but that head noun animacy is not the only motivation for passive choices in Mandarin. Thus we hypothesize that when both nouns are animate, participants will be more likely to produce passives in which they omit the agent compared to conditions with one animate and one inanimate noun (Gennari et al., 2012).

Methods

Participants

Thirty-two native speakers of Mandarin Chinese participated in the experiment for course credit in an introductory psychology class at University of Wisconsin-Madison. For this and all experiments reported here, all participants indicated that they had been born and educated in Mainland China or Taiwan and spoke Mandarin Chinese as their dominant language. Most were freshmen or sophomores who did not arrive in the US until the start of their university attendance. Participants took part in no other experiments reported in this article. All phases of the experiment, including introductory description of the study, obtaining informed consent, and instructions for the task, were conducted in Mandarin with a native Mandarin-speaking experimenter.

Materials and design

Twenty pictures were used as the experimental stimuli. Each picture showed two human agents performing the same action. One of them acted upon an animate patient (e.g. kissing a girl) and the other acted upon an inanimate theme (e.g. kissing a trophy). Also in each picture were additional human characters and objects that were similar to the patient/theme, such as one or more additional girls and trophies in the kissing picture. These similar items were designed to encourage spoken responses with noun modification, especially relative clause modifiers (e.g.

"the trophy" doesn't distinguish among the two trophies in a picture, and so speakers need to add additional modification to clarify which trophy they are talking about). See Fig. 1 for an example experimental picture. Thirty filler pictures were also created. They mostly depicted a scene where a single main event was happening; for example, some people bowling in a bowling alley or several people playing poker.

A native Mandarin speaker recorded two audio questions for each experimental picture. One of the questions targeted the animate patient and the other the inanimate object. For example, the questions recorded for the picture in Fig. 1 were "Who has black hair?" and "What is yellow?", for the animate and inanimate patient/theme of *kiss*, respectively. The structure of the questions and the nature of the pictures (i.e. more than more one girl and one trophy) together encouraged relative clause responses, which specifically distinguish one entity from others (i.e., animate: *the girl that a woman is kissing/the girl that is being kissed by a woman*; inanimate: *the trophy that a girl is kissing/the trophy that is being kissed by a girl*).

One question was recorded for each filler item. Filler questions contained a variety of structures and were not designed to elicit relative clauses. Two scripts were prepared to counterbalance the assignment of questions to experimental pictures, so that each participant heard ten animate-targeting and ten inanimate-targeting questions, but was asked only one question for each picture. In each script, experimental trials were interleaved with fillers so that no more than two experimental pictures occurred in a row.

Procedure

Pre-training task. A pre-training task was administered before the main production experiment. Participants were shown pictures depicting individual actions that would appear later in the actual task. These pictures were

Fig. 1. An example of the experimental item for the production task in Experiment 1. The target items were the girl and the trophy being kissed. Pictures contained additional entities that resembled the target (e.g. two trophies, one yellow and one orange), making it felicitous to use relative clauses to answer the question (e.g. what is yellow?). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 2. An example picture that described the action "kiss" in the pre-training task.

cropped from the experimental items. See Fig. 2 for an example. Two seconds after the presentation of the picture, the verb that described the action (i.e. *qin* kiss) appeared beneath the picture. Participants were asked to read the verb aloud. The purpose of the pre-training task was to ensure that participants could easily identify the pictured action and would use the designated verb in their production in the actual test so that responses were comparable across participants. Each participant viewed two pictures associated with a verb: the action performed on the animate entity and the inanimate object. The filler pictures were also cropped from the original filler pictures but the accompanying words were nouns instead of verbs. Inclusion of fillers and both animate and inanimate targets in the pre-training was designed to avoid calling attention to the experimental items in the main task.

Main task. Upon completion of the pre-training task, participants received instructions about the main task. Participants were told that they would see pictures and answer questions about them. To prompt production of relative clauses, the participants were told that there were multiple similar characters and objects in the pictures and their answers would be shown to a later group of people to guide their recognition of specific characters or objects in the pictures. Participants heard a cover story indicating that elements in terms of color or screen position would change in the pictures presented to the later group and but the actions that the pictured entities were involved in wouldn't. Because modification with an action description is accomplished with relative clauses, participants were therefore encouraged to produce relative clauses without having these structures explicitly mentioned in the instructions.

The task began with three practice trials. At the beginning of each trial, a color picture like the one in Fig. 1 appeared. Following a three-second delay allowing the participants to inspect the picture, a pre-recorded question was played through speakers. Participants' spoken response to the question was digitally recorded for later

analysis. The picture remained on the screen during the spoken response and was removed by a keypress after the participant finished answering. A new trial began immediately afterward.

Coding and results

All responses that did not contain relative clauses were excluded. This process removed 16% of animate trials and 21% of inanimate trials. This level of non-relative clause responses is similar to that in previous studies (Gennari et al., 2012; Montag & MacDonald, 2014) and is not surprising, as participants were never explicitly instructed to produce relative clauses. Excluded trials included ones in which participants did not notice the additional characters or objects that resembled the target and answered only with a noun (e.g. trophy), and ones in which they violated the instructions to avoid color or positional information, producing responses such as *the yellow trophy*, or *the trophy on the right*.

The participants' relative clause responses were coded as Passive or Active. Analyses in all experiments here were conducted with linear mixed-effects models, using the function lmer in the lme4 package in the statistical software R. A logit mixed effects model was fitted for the data with Animacy as the fixed factor and Subjects and Items as random factors with random intercepts and slopes. As in languages with head-first relative clauses (English, Spanish, and Serbain; Gennari et al., 2012; Montag & MacDonald, 2014), animacy significantly influenced participants'choice of relative clause structures ($\beta = -0.243$, $SE = 0.066$, $t = -3.68$). An overwhelming 98% of responses describing animate entities were passives, while 75% of responses in the inanimate condition were passives, shown in Fig. 3.

We also examined, within passive responses, the rate at which participants omitted the agent of the action in bare passives like *[bei kiss de] trophy* vs. full passives with the agent noun, e.g. *[bei girl kiss de] trophy*. Animacy of the head noun significantly affected whether the agent nouns
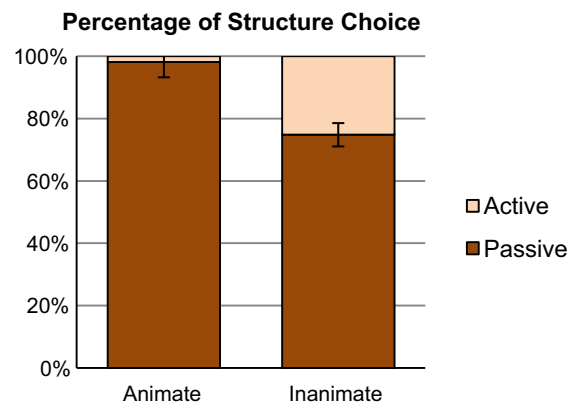

Fig. 3. The percentage of structural choice as either in the active or passive relative clause forms when describing animate or inanimate entities. The error bars indicate standard error of means.
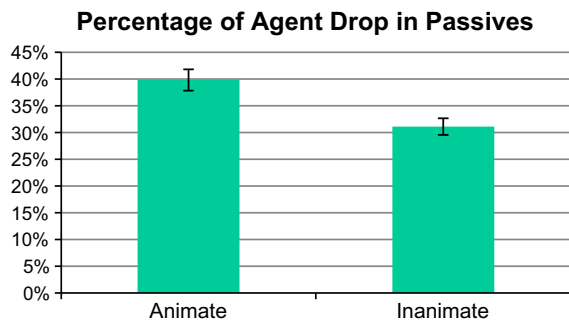
Percentage of Agent Drop in Passives

Fig. 4. Percentage of agent-drop among passive relative clause responses as a function of animacy of the head noun. The error bars indicate standard error of means.

were omitted ($\beta = -0.10$, $SE = 0.04$, $t = -2.67$). 40% of agents were dropped in the animate condition compared with 31% in the inanimate condition, as shown in Fig. 4. This result replicates findings in English and Spanish (Gennari et al., 2012; Montag & MacDonald, 2014).

Discussion

In this picture description task, Mandarin speakers showed an overwhelming passive preference in describing an animate entity, comparable to data from English speakers (Gennari et al., 2012; Montag & MacDonald, 2014), and higher than the rates for Spanish and Serbian speakers reported in Gennari et al. The 75% passive rate for inanimates was also very high; English, Spanish, and Serbian speakers in a similar task produced 50% or fewer passives (Gennari et al., 2012; Montag & MacDonald, 2014). Thus Mandarin speakers find passive relative clauses a very viable option for describing both animate and inanimate entities in this spoken production task.

While this study did not investigate factors (other than animacy) that could affect passive rates, we can speculate why passive is the dominant choice in this study. High rates of passives may be attributable to an interaction among demands of on-line spoken production, the relative clause word order in Mandarin, and noun animacy. To passivize a relative clause in Mandarin, one only needs to add a passivizer *bei* at the beginning of the phrase. Thus the choice to passivize must be made early, but the realization should be quite easy, involving only a single high-frequency morpheme. The early position of *bei* and its optional use may allow it to have a delaying function, similar to the relative pronoun and compelementizer *that* function in English—speakers use these forms in part to provide more time to plan their utterance, especially when they encountered production difficulty (Ferreira & Firato, 2002; Jaeger, 2005; Jaeger & Wasow, 2005; Race & MacDonald, 2003). We are not claiming that added planning time from the addition of phrase-initial *bei* is the sole reason of passive use in Mandarin, but future work should investigate whether this feature of passive formation in Mandarin contributes to high rates of passive here.

Another incentive for passivization in Mandarin is that while both agents and patients must be mentioned in object relatives, as in English and many other languages, the agent of the action can be omitted in passives. Gennari et al. (2012) argued that the option to omit the agent was a force behind some passive use, as agent omission is beneficial under conditions of similarity-based interference, where both the agent and patient are semantically similar (e.g. *girl* and *woman*). The Mandarin results are consistent with these claims for relative clauses and also consistent with results in simple main clauses in Mandarin, in which agent omission is higher under conditions of similarity-based interference (Hsiao, Gao, & MacDonald, 2014). Again, the claim is not that similarity-based interference is the only motivation for agent omission, but rather that it is a reliable contributing factor.

The strong effects of animacy in this study are notable because of the special features of Mandarin relative clauses. First, owing to the head-final word order of Mandarin relative clauses, the animacy effects show that a late word (the head noun) affects the earliest word in the relative clause. This result argues against the most radical accounts of incremental processing (De Smedt, 1996; Kempen & Hoenkamp, 1987; Levelt, 1989) because it suggests that speakers must have planned at least some aspects of the head noun before uttering the first words of the relative clause. Second, as we've noted, the active and passive forms are identical in words and word order except for the use of *bei*. Thus in contrast to many other languages, the active/passive choice here has minimal consequences for the surface form of the utterance, and yet speakers are strongly affected by head noun animacy in their choice of these two forms.

We next investigate the consequences of these relationships for the distributional statistics of the language, using a corpus of written Mandarin. We chose a written corpus because in many languages, the use of passives can differ in written and spoken genres, including active and passive relative clauses in English (Montag & MacDonald, 2015; Roland, Dick, & Elman, 2007). A corpus analysis also allows us to look beyond the particular materials of our production study. The pictures we used in the production task always involved transitive physical actions that often (though not always) implied adverse outcomes or infliction, which is the connotation embedded in the Mandarin passive morphemes, including *bei*. Analysis of a written corpus would also allow us to examine the effect of the specific discourse context created for the current production task. Since our test of comprehension will be reading times, to be maximally comparable to prior relative clause comprehension studies (Gibson & Wu, 2013; Hsiao & Gibson, 2003; Lin & Bever, 2006; Vasishth et al., 2013; Wu et al., 2012), we will use the written corpus to investigate passive and active relative clause use in texts.

Corpus analysis

The aim of this study was to investigate the relationship between head noun animacy and relative clause structure in Mandarin. While there are other previous corpus

Table 1
Frequency counts of object relatives and passive relatives in the Chinese Penn Treebank 7.0.

	Animate head noun	Inanimate head noun	Total
Object relatives	21	288	309
Full passive relatives	10	7	17
Bare passive relatives	10	14	24
Total	41	309	350

analyses of Mandarin relative clauses (Hsiao & Gibson, 2003; Hsiao & MacDonald, 2013; Jäger et al., 2015; Lin, 2011; Pu, 2007; Wu, 2009), those analyses did not fully considered these lexico-syntactic pairings. Wu (2009), Lin (2011) and Hsiao and MacDonald (2013) all conducted corpus analyses on Mandarin relative clauses that included noun animacy as a factor, but none fully investigated the conditions relevant to the production data in Experiment 1 (Lin, 2011, investigated animacy in passive relative clauses in general, without distinguishing full and bare passives). The strong animacy effects observed in Experiment 1 suggest that corpus analyses considering only overall frequencies of relative clause types will not adequately capture the important animacy–structure correlations that comprehenders are likely to use in comprehension. That is, even though the animate head noun follows the relative clause, we expect that there are correlations between animacy and structure type that help comprehenders interpret relative clauses and anticipate the animacy of an upcoming head noun.

Methods

Hsiao and MacDonald (2013) investigated several kinds of active relative clauses in the Chinese Penn Treebank 7.0 (Xue et al., 2010), but not passive relatives. We built on their coding and extracted passive relative clauses from the same source. The Chinese Penn Treebank is a parsed corpus consisting of more than one million words, with various sources from the news genre, such as newswire data, broadcast materials, web texts, magazine articles, from Mainland China and Hong Kong. We used TGrep2 1.15 (Rohde, 2005) to extract passive relative clauses, including both full passives (with agent) and agentless forms (see Appendix A for search patterns). We hand coded the head noun animacy in each relative clause, following the criteria used by Hsiao and MacDonald (2013): animate entities refer to living beings with agency and volition to perform an action, whereas inanimate entities refer to those without this property (Hundt, 2004). Coding was performed by a Mandarin-speaking research assistant.

Results

The frequencies of object and passive relatives coded for the animacy of their head nouns are reported in Table 1. The first row of the table indicates that object relatives are much more frequent with inanimate head nouns than animate ones, a result that is consistent with prior corpus results in English (Gennari & MacDonald, 2009).

If comprehenders are sensitive to these contingencies, then they should tend to expect inanimate heads for object relative clauses, assuming that the lexical content makes this plausible, and also assuming that comprehenders know that they are encountering an object relative.

The table also shows striking differences to the picture description responses in Exp.1, with far fewer passives in the written texts than in our spoken production task (given the low counts of passives in the corpus, we cannot evaluate the animacy effect here). Lin's (2011) data, also based on a written corpus, also showed that the frequency of passives is considerably lower than the typical object relatives (about 1:9) and there is not a clear bias between animate and inanimate head distribution. Roland et al. (2007) also reported a frequency disparity of passive relatives between English speech and text, but in the opposite direction: passives are rarer in the spoken than in the written corpora. They attribute the difference to discourse register: one of the functions of passives is to de-emphasize the agent, which is less necessary in conversation when the speakers and the speakers' actions are the focus of discussion. Discourse register may also be a plausible cause for the difference found in speech and text in Mandarin but with different language-specific factors. Passive markers, including *bei* and other rarer markers, tend to have a negative connotation related to infliction (Chao, 1968). In their corpus analysis comparing different genres of language use, Xiao, McEnery, and Qian (2006) found that passives were least frequent in newspaper, news editorials and official documents. Because the Chinese Treebank consists of texts from news sources, the low rate of passives in the Chinese Treebank may be due to a tradition of low passive use in this genre. Xiao et al. suggested that the high rate of passives in English written texts, especially in formal documents, is due to the function of passives to mark objectivity, whereas passives in Chinese do not have this function. They further noted that the passive marker *bei* is used to denote a predominantly negative tone in speech but usually not in official documents. This potentially accounts for the high rate of passivization in our production data in Experiment 1, where the materials often involved physical infliction and damage.

These results suggest that genre has a large effect on passive use in Mandarin. The spoken production task in Experiment 1 appears to have created conditions (adverse events) for which passives are highly felicitous, whereas the Chinese Treebank appears to have selected genres in which passive use is quite rare. Both the spoken production data and the corpus data point to important correlations between head noun animacy and relative clause structure, but neither of these studies assessed passive use in a discourse environment that would be useful for a comprehension experiment—written materials for which passive relative clauses would be a viable option. We therefore developed new stimuli for a written sentence comprehension study. These stimuli crossed relative clause type (object relative, passive relative) and head noun animacy, including a preceding context to make the sentence plausible. We examined interpretations of these stimuli first in a gated sentence completion task in Experiment 2 and then in a self-paced reading task in Experiment 3.

Table 2
Example items for the sentence completion task.

Condition			Example
Animate head-oriented context 因为不愿意引起轩然大波, Unwilling to cause an uproar	Early Gate	Active	媒体 批评＿＿＿＿＿＿ 。 media criticize＿＿＿＿＿.
		Passive	被 媒体 批评＿＿＿＿＿＿ 。 *bei* media criticize＿＿＿＿＿.
	Late Gate	Active	媒体 批评 的＿＿＿＿＿ 。 media criticize *de* ＿＿＿＿.
		Passive	被 媒体 批评 的＿＿＿＿＿。 *bei* media criticize *de* ＿＿＿.
Inanimate head-oriented context 虽然剧情沉闷了无新意, Although the plot was dull and unoriginal,	Early Gate	Active	媒体 批评＿＿＿＿＿＿ 。 media criticize＿＿＿＿＿.
		Passive	被 媒体 批评 ＿＿＿＿＿＿ 。 *bei* media criticize＿＿＿＿.
	Late Gate	Active	媒体 批评 的＿＿＿＿＿ 。 media criticize *de* ＿＿＿＿.
		Passive	被 媒体 批评 的＿＿＿＿＿。 *bei* media criticize *de* ＿＿＿.

Note: Early vs. Late Gate was manipulated between subjects, whereas structure type and context type were manipulated within-subjects.

Experiment 2 – gated sentence completion task

In this experiment, we investigated comprehenders' syntactic and semantic interpretations at various points in Mandarin relative clauses using a gated sentence completion task (McRae et al., 1998) with relative clauses. Using this method, Gennari and MacDonald (2008) found that the animacy of the head noun affected the degree to which participants adopted an object relative clause interpretation of the subsequent words, which in turn predicted comprehenders' reading times for sentences that turned out to be object relative clauses. Our investigation here is similar, but the differences in word order between English and Mandarin do not allow us to examine exactly the same factors. English relative clauses are head-first, permitting the study of early-arriving head noun properties such as animacy on the subsequent relative clause. Mandarin has head-final relative clauses, and so head noun animacy is not a predictor but potentially something that is predicted by other, earlier linguistic material.

Another consequence of the head-final relative clause structure in Mandarin is ambiguity. Comprehenders do not necessarily know that the structure they are perceiving is a relative clause. In particular, simple sentence structures are a much more common alternative interpretation for the Noun + Verb word order that initiates an object relative clause (see Table 2 for examples of active fragments) (Hsiao & MacDonald, 2013). The relativizer *de* provides strong support for a relative clause interpretation (though other interpretations of *de* exist, see Hsiao & MacDonald, 2013, for details), and our choice of gates was designed to identify the interpretations that dominate before and after encountering *de*. We investigated whether comprehenders consider a relative clause interpretation of the input before the relativizer *de*, and whether the sentence initial passive marker *bei* influences expectations for a relative clause. *bei* may promote a relative clause interpretation because it functions as a light verb, and a sentence starting with a verb is uncommon, as it is only seen in pro-drop sentences, gerunds or subject relatives, including passive relative clauses. At the point where *de* makes a relative clause highly likely, we investigated whether completions favor an animate or inanimate head noun. Experiment 1's production data and corpus results both showed that object relatives are more common with inanimate head nouns, whereas passives were more common with animate heads in Experiment 1, and we expect these relationships between active/passive and head animacy will arise in our completions.

Methods

Participants

Fifty-six native speakers of Mandarin Chinese on the campus of University of Wisconsin-Madison participated in the experiment. All of them read and wrote in Simplified Chinese, which was the written script adopted in this experiment.

Materials and design

We developed context + relative clause fragments for the sentence completion task, where the expected completion, if participants interpreted the fragment as a relative clause, would be a head noun for the relative clause, followed by a main clause with the head noun as the main clause subject. Because relative clauses are much more felicitous in some discourse contexts than others (in Mandarin and in other languages), we developed an introductory context phrase that made additional relative clause modification felicitous. As it is effectively impossible to develop a discourse context that is equally appropriate for an animate and an inanimate head noun modified by a relative clause, we developed two contexts for each fragment, one that supported relative clause modification of an animate head, and one that supported a relative clause modification of an inanimate head. Examples are shown in Table 2.

Because the relative clause head serves as the main clause subject in a completed sentence, the inanimate head context necessarily promotes an inanimate main clause subject. Inanimate subjects are rarer than animate ones in the Mandarin Treebank corpus (except in the case of financial reports such as *The economy grew*, Hsiao & MacDonald, 2013), and so this context necessarily promotes a comparatively rare sentence type. The animate-promoting context

promotes a more typical animate main clause subject (the relative clause head), but because the relative clause noun precedes the head noun, it is likely that this first noun may be interpreted as the main clause subject rather than part of the relative clause. As we have noted, this is the essential ambiguity of the head-final Mandarin relative clauses. A key question is whether the passive marker *bei* can reduce this ambiguity effect.

To investigate this question, the potential relative clause material in the fragment was presented either in active form or in passive form, via the insertion of the passive marker *bei* clause-initially. The fragments were truncated at two different gates for completions. At the Early Gate, fragments had a "(*bei*) Noun Verb" sequence and ended before the appearance of a relativizer *de*. The Late Gate was identical but contained one more word, the relativizer *de*, which strongly promotes a relative clause. As Table 2 shows, these manipulations yielded a $2 \times 2 \times 2$ design, with Context Type (animate head-oriented vs. inanimate head-oriented) and Structure Type (active vs. passive) as within-subject factors and Gate (Early vs. Late) as a between-subject factor. Gate was manipulated between-subjects so as not to introduce a priming effect, as participants exposed to both Early and Late Gates could potentially be primed to complete the Early Gate fragments with relative clauses from having encountered many fragments ending with the relativizer *de*. A total of 28 items were created, counterbalanced in eight scripts (see Appendix B for the full list of experimental items).

Procedure

The experiment was conducted entirely in Mandarin with a native-speaking experimenter. Participants were randomly assigned to the Early or Late Gate condition and were given a paper survey that listed the sentences to be completed. The active and passive sentence fragments were semi-randomly ordered so that no more than three of the same sentence type (i.e. by context type and by active/passive voice) appeared in a row. The same randomized order was applied to each script and to every participant (i.e. all eight scripts had the same item order and all participants who were assigned to a certain script received the same order of stimuli). No time limit or minimal length of a sentence was imposed. Participants were encouraged to continue the sentences until they felt the sentences were complete in the semantic and grammatical sense, and they were told that there were no right answers. Participants' completions were coded by research assistants who were native speakers of Mandarin. Completions were coded for whether they adopted a relative clause interpretation or some other interpretation of the (*bei*) + Noun + Verb sequence, and in the case of relative clause interpretations, whether they were completed with an animate or inanimate head noun. The codings were double checked by a second coder to ensure accuracy.

Results

The results at the Early Gate and the Late Gate, given either the object relative or passive fragments, separated by context type, are presented in Fig. 5.

Early Gate

Without the relativizer *de* at this gate, completions indicate whether alternative interpretations existed before a clear relative clause cue appeared. Across the two context types, there was a large effect of structure on participants' completions: for passive fragments, 51% of the completions were relative clauses compared to only 13% in the active condition ($\beta = 0.40$, $SE = 0.03$, $t = 15.24$). Most continuations for the active fragments were a simple sentential object, making the sentence a simple Subject–Verb–Object sentence. The high proportion of relative clause completions in the passive condition suggested that the passive marker *bei* provided a strong incentive for participants to continue the sentence with a relative clause structure. Among all the valid active relative clause completions, around 41% were completed with an animate head noun, compared to 60% for passive relative clause completions ($\beta = -0.15$, $SE = 0.07$, $t = -2.01$). The context phrases appeared to constrain the animacy of the head nouns for those completions with relative clauses. Among the very few relative clause completions for object relative fragments, the animate head-oriented context elicited completions which an animate head 70% of the time. The inanimate head-oriented context elicited 82% of inanimate heads for completions with a relative clause. For passive fragments, the pattern was 89% animate heads in the animate context and 68% inanimate heads for the inanimate context. The relatively higher constraining power of the animate context (and lower constraint of the inanimate context) for the passives seems to reflect the preference of using passives to modify animate nouns ($\beta = -0.20$, $SE = 0.06$, $t = -3.64$), consistent with the results of Experiment 1.

Late Gate

At this gate, the relativizer *de* was provided in the sentence fragment, yielding 89% relative clause completions in the active frame and 97% in the passive frame ($\beta = 0.06$, $SE = 0.02$, $t = 4.12$). The non-relative completions reflected other interpretations of *de*. 62% of the relative clauses in the passive condition were completed with animate head nouns, whereas 46% had animate head completions in the active condition ($\beta = -0.16$, $SE = 0.03$, $t = -4.59$). In the active frame, the animate head-oriented context elicited 67% of relative clause completions with animate heads, whereas inanimate context yielded 75% of inanimate heads completions. For the passive frame, the percentage was 91% for animate context and 68% for inanimate context. These patterns resemble those of the Early Gate. The passive fragments produced more conforming responses corresponding to context phrases targeted toward animate nouns ($\beta = -0.24$, $SE = 0.04$, $t = -6.32$).

Discussion

This study yielded several important results. First, the completions at the Early Gate are a testament to the extensive ambiguity of relative clauses in Mandarin, with most participants interpreting the relative clause fragments as main clauses. This was the case for both the animate and inanimate context conditions. Second, the difference between the passive and object relative fragments showed
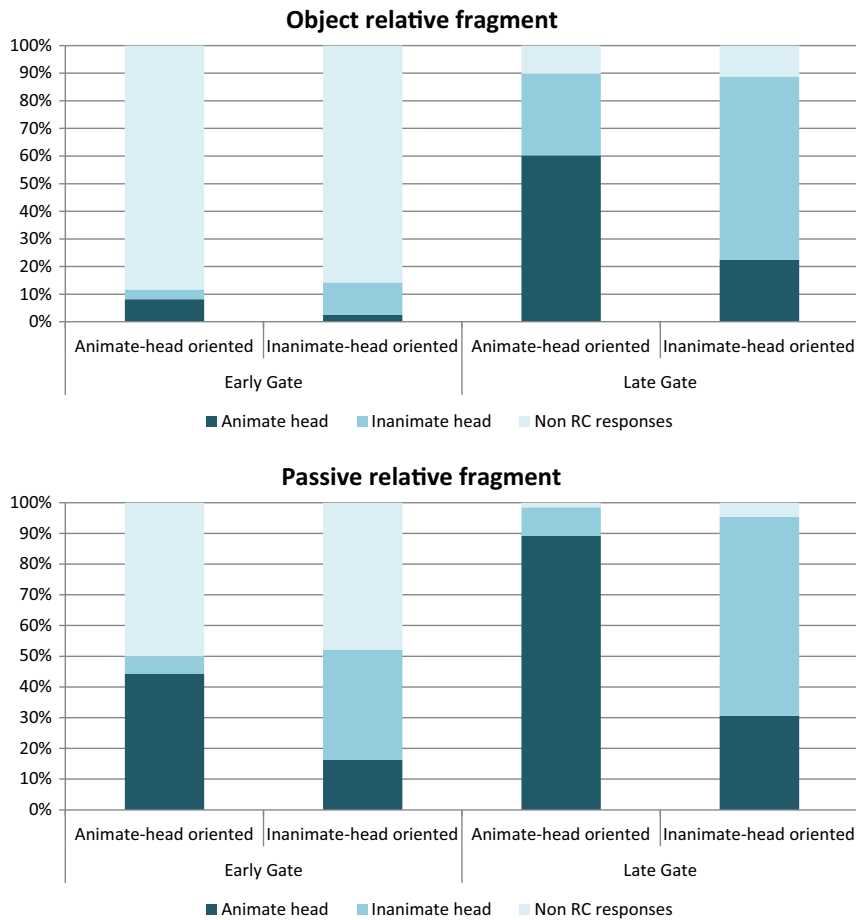
Object relative fragment

Passive relative fragment

Fig. 5. The two panels display the percentage of completions proceeding either an active object relative fragment or a passive relative fragment. In each panel, the different colors in each bar represent types of completions distinguished as (1) relative clauses with animate heads, (2) relative clauses with inanimate heads, and (3) non-relative clauses. The responses vary on the dimensions of context type (animate head-oriented, inanimate head-oriented) and Gate (Early Gate, Late Gate).

that, even though relative clauses are head-final and the (largely) disambiguating relativizer arrives late, the early portions of a relative clause affected comprehenders' expectations for upcoming material. In particular, the result at the Early Gate showed that the passive marker *bei* was a strong cue for a relative clause even before the relativizer *de*. This result is important because main clause completions are fully grammatical at the early gate, and relative clauses are much less frequent than main clauses, and yet, in the presence of *bei*, there were nonetheless a number of relative clause completions. This result suggests that in an appropriate discourse context in written text, passive relative clauses are a viable option for Mandarin speakers. Third, the presence of *bei* not only promoted a relative clause interpretation, but it more particularly led participants to expect a relative clause that modified an animate head noun. This result suggests that comprehenders generate expectations about the pairing of passives and animate heads, consistent with the patterns in the production study. Finally, these results showed the degree to which our contexts affected interpretations: before *de*, relative clause completions were extremely rare, showing the dominance of simple main clause interpretations at this

point, independent of the context manipulation. At the late gate, however, our contexts affected animacy in the completions: animate-promoting contexts did promote more animate head completions than inanimate ones, while inanimate-promoting contexts promoted more inanimate head completions.

The next experiment investigates the consequences of these expectations, using the Experiment 2 materials, expanded to be complete sentences, in a comprehension task. Common animacy–structure pairings should result in easier comprehension, as measured by reading times, than uncommon pairings, as in Gennari and MacDonald (2008, 2009). Alternatively, if comprehension difficulty is affected by dependency distance (Gibson, 1998, 2000), then passive relatives, which have a longer dependency distance, should be harder than object relatives.

Experiment 3 – comprehension task

In this experiment, we investigate the relation between the interpretation preferences of the gated sentence completions in Experiment 2 and the reading times in sentence

comprehension. We hypothesized that the degree of ambiguity revealed in the completion study would positively correlate with reading times. For example, we expect that the high rate of passive relative clause completions compared to the active completion rate in Exp. 2 at the Early Gate would predict shorter reading times for the early part of passive relative clauses, and that the high rate of animate head continuations after the passive relative fragments at Late Gate would predict shorter reading times for passive relative clauses after the head nouns were revealed to be animate.

Methods

Participants

Forty native speakers of Mandarin Chinese at the University of Wisconsin-Madison participated. All of them read Simplified Chinese. Ten participants were assigned to each of the four counterbalanced scripts.

Materials and design

Two factors were manipulated: the structure of the relative clause (active, passive) and the animacy of the head noun (animate, inanimate), yielding a 2×2 within-subject experimental design. As in Experiment 2, each stimulus sentence began with a context phrase that created an expectation for an upcoming noun phrase and its corresponding animacy properties. Following the context, the critical sentence always had the sequence: [relative clause] + head-noun + adverb + matrix verb..., followed by one or more additional words to make a sensible sentence. The relative clause always modified the main clause subject, and this subject (head) noun was always immediately followed by an adverb, then the main clause verb. The purpose of including an adverb was to lengthen the number of words that were comparable across items. Examples of the four conditions are in Table 3. As these examples show, the only difference between the active and passive relative clauses was the presence/absence of the passive marker *bei*.

There were 28 experimental items in total, 7 in each condition. 50 filler sentences were also created, containing various syntactic structures. Four scripts were created to counterbalance the assignment of items to conditions. A comprehension question was constructed for each test and filler sentence. Half of the questions for the experimental sentences were focused on the head noun and the other half on the relative clause-internal noun. The distribution of correct answers was also counterbalanced across both test and filler items (see Appendix B for the full list).

Procedure

Sentences were presented with Linger software (Rohde, 2001), which allowed self-paced word-by-word reading with a moving window display. The stimuli were presented in a quasi-random order so that no more than two experimental items appeared adjacent to one another. Each participant received a different randomized order.

Participants were instructed to read for comprehension and answer a question after each sentence. Feedback was given for incorrect responses to the comprehension question, whereas correct answers would trigger the next trial with no feedback. Six practice trials were administered before the actual experiment. Participants could choose to take a short break after finishing half of the experiment.

Results and discussion

Mixed effects models were applied for statistical analysis, with Structure and Animacy being fixed factors and subjects and items being modeled for random intercepts. The R package *lme4* was used for modeling and the library *languageR* was used for significance testing.

Comprehension question accuracy

The mean comprehension accuracy was high overall: 93.93% for experimental items and 91.21% for fillers. There was no difference in accuracy and comprehension question response times among the four conditions (animate active: 95.36%, 1802 ms; inanimate active: 91.58%, 1713 ms; animate passive: 96.42%, 1607 ms; inanimate passive: 92.36%, 1900 ms), suggesting that participants' attention and item difficulty were comparable across conditions. Due to the high overall accuracy, no data were removed.

Word-by-word reading times

The analysis was conducted on the reading times in a seven-word critical region: *bei* (present only in the passive

Table 3
Example items in the comprehension task.

Condition	Example
Animate-headed active	因为 不愿意 引起 轩然大波, [媒体 批评 的] 歌手 低调地 不 做 任何 回应。 Because unwilling cause uproar, [media criticize *de*] singer low-key NEG make any comment Unwilling to cause an uproar, the singer that the media criticized made no comments in a low-key way.
Animate-headed passive	因为 不愿意 引起 轩然大波, [被 媒体 批评 的] 歌手 低调地 不 做 任何 回应。 Because unwilling cause uproar, [*bei* media criticize *de*] singer low-key NEG make any comment Unwilling to cause an uproar, the singer that was criticized by the media offered no comments in a low-key way.
Inanimate-headed active	虽然 剧情 沉闷 了无新意, [媒体 批评 的] 电影 意外地 获得 大奖。 Although plot dull unoriginal, [media criticize *de*] movie surprisingly won major award Although the plot was dull and unoriginal, the movie that the media criticized surprisingly won a major award.
Inanimate-headed passive	虽然 剧情 沉闷 了无新意, [被 媒体 批评 的] 电影 意外地 获得 大奖。 Although plot dull unoriginal, [*bei* media criticize *de*] movie surprisingly won major award Although the plot was dull and unoriginal, the movie that was criticized by the media surprisingly won a major award.

Note: The sentences in Mandarin are marked with word boundaries by spaces here. Word-by-word glosses are provided here, along with full-sentence translation at the bottom in each cell. Relative clauses are marked with brackets here, but brackets were not shown to participants.

condition), the relative clause subject (n1), the relative clause verb (rv), *de*, the head noun (n2), the post-head adverb (adv) and the matrix verb (mv).

Reading times, including those of the filler trials, underwent trimming if beyond 2.5 standard deviation of the participant's mean or longer than 4000 ms, affecting 2.6% of all data points. These reading time data were analyzed following the steps for a two-stage data treatment described in Jaeger's Lab Blog (http://hlplab.wordpress.com/2008/01/23/modeling-self-paced-reading-data-effects-of-word-length-word-position-spill-over-etc/). First, a log transformation was applied to normalize the distribution. The log-transformed reading times were further regressed on a series of predictors to control for extraneous effects, with random intercepts for each participant: difference between fillers and test items ($\beta = -0.016$, $SE = 0.003$, $t = -5.57$), word length ($\beta = 0.052$, $SE = 0.002$, $t = 26.70$), the log-transformed position of the trial in the list ($\beta = -0.108$, $SE = 0.002$, $t = -70.23$), and the restricted cubic spline of the word position in the sentence (we followed the suggestion of one reviewer by equating the position of *bei* with that of the following word in order to avoid stripping away the potential effect of filler-gap distance) ($\beta_0 = 0.016$, $SE = 0.003$, $t = 5.43$; $\beta_1 = -0.040$, $SE = 0.017$, $t = -2.04$; $\beta_2 = 0.030$, $SE = 0.066$, $t = 0.45$; $\beta_3 = 0.145$, $SE = 0.077$, $t = 1.89$). The residual log reading times were computed from the above model and served as the dependent variable for the analyses in the next phase. Fig. 6 shows on the two panels, separated by animacy, the mean residual log reading times of active and passive relatives at each word position in the 7-word region.

The factors of interest (Animacy, Structure) and their interaction were applied to the second phase of the analyses. Even though the animate vs. inanimate head noun contrast does not appear until word 5, Head noun animacy is included as a factor throughout the analyses because the preceding discourse context varied in setting up a plausible animate or inanimate head noun. Because the word *bei* only existed in the passive condition, no structure effect was modeled at this position. Spillover effects of the three words preceding a given word *n* were also included in all

analyses: *n*-1, *n*-2, and *n*-3. Here we describe the effects of Animacy, Structure and their interaction; the full details of statistical results, including *p* values, can be found in Appendix C.

Early word positions. A significant animacy effect was found at *bei* ($\beta = 0.055$, $SE = 0.023$, $t = 2.36$), reflecting the association of animacy and passive use: reading times for *bei* were shorter following the animate-head promoting context than in the inanimate context condition. The reading times for the next four words (up to the head noun) did not show any main effects of animacy, structure, or the interaction. However, a simple comparison within each level of animacy indicated significant difference in reading times between object and passive relatives at the relative clause subject (n1) position: passives were read faster than actives in the inanimate condition ($\beta = -0.054$, $SE = 0.023$, $t = -2.41$) but not in the animate condition. A numerical but non-significant difference was observed between animate-headed object relatives and passive relatives ($p = .13$).

Later word positions. Fig. 6 shows that at the adverb and the main verb, one condition – the object relatives with animate heads – is more difficult than the other three. Specifically, at the adverb position, there was a reliable interaction of animacy and structure ($\beta = 0.077$, $SE = 0.031$, $t = 2.52$), such that animate-headed object relatives were read reliably more slowly than animate-headed passives ($\beta = -0.062$, $SE = 0.022$, $t = -2.82$), and no difference at this position for the inanimate-headed sentences. Similar patterns were found in the last word region, the main verb (mv). There was again a reliable interaction ($\beta = 0.097$, $SE = 0.033$, $t = 2.97$), such that object relative sentences with animate heads were read significantly more slowly than animate-headed passive relatives ($\beta = -0.115$, $SE = 0.024$, $t = -4.84$).

In summary, both animacy and structure had clear effects on comprehension difficulty. Beginning with a marginal effect at the head noun and continuing with reliable effects in the following two words, active object relative
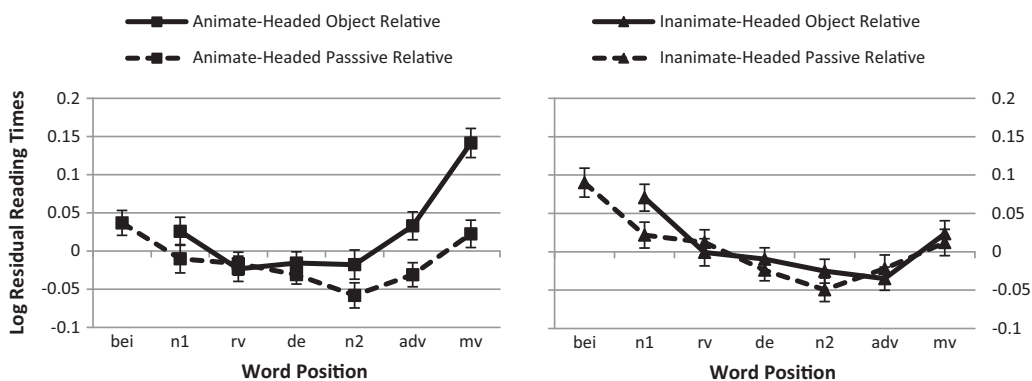

Fig. 6. The mean log residual reading times of active and passive relative clauses at each word position, with the left panel showing animate-headed relative clauses and the right panel inanimate-headed relative clauses. *bei* = passive marker (in passive condition only), n1 = first noun of relative clause (the agent of the action), rv = relative clause verb, *de* = relativizer, n2 = head noun, adv = adverb, mv = main verb.

sentences with animate heads were the hardest among the four conditions. This particular pattern may owe in part to the context promoting an animate subject (the relative clause head), potentially contributing to the misinterpretation of the relative clause noun. However, since animate subjects are very common, this context reflects typical patterns in the language, and the misinterpretation here also reflects the basic ambiguity in Mandarin relative clauses, where the relative clause noun is misinterpreted as a main clause noun, especially when this noun is animate. Importantly, the pattern was very different in the passive condition: by simply adding a *bei* in front of this embedded noun, people quickly settled on a relative clause reading. The results suggest that the passive marker significantly reduced ambiguity and provided a strong cue, especially when the modified nouns were animate, matching what we found in the previous two experiments. Similar interactions between structure and animacy have been observed in English, in spite of cross-linguistic differences in relative clause head direction. Gennari and MacDonald (2008, 2009) found in their self-paced reading data that active relative clauses with animate heads were also the hardest to comprehend. They used regression analyses to link expectations, as measured by their completion studies, to reading times, and we pursue similar analyses below.

Regression of sentence completion data onto reading times

To investigate the effect of expectations (completions) on reading times, we conducted regression analyses to relate the mean completion patterns for each item to the mean reading times for that item. As noted in Experiment 2, the presence of *de* can strongly alter the interpretation of the sentence, especially the active object relative clause structure, because those sentences are consistent with a simple main clause interpretation before *de* is encountered. Because there were very few relative clause completions at the Early Gate, we did not include expectations for a particular head noun animacy at this position and instead regressed only the average rate of relative clause completions for each item. At the Late Gate, after *de* had been encountered, relative clause completions predominated, and at this position we used as a predictor the average rate of relative clause completions with the matching head noun animacy specified by the context phrase (identical to the head noun animacy of the expanded sentences used in the comprehension study).

Our original plan was to regress the completion data for each item on critical word reading times for this item. We did not initially include the presence/absence of the passive marker *bei* as a factor in the regression, because both completion and reading participants saw *bei* at the start of the relative passive relative clauses, and so we assumed that this passive/active context information would be reflected in completions. Similarly, we did not initially include an Animacy factor, since the major animacy manipulation, the early context that promotes an animate or inanimate head noun, is read by both completion and reading participants, and thus the Animacy context should be captured in the completions. Our initial regressions did find reliable effects of completions on reading times (the

Early Gate completions reliably predicted reading times at the head noun and the Late Gate predicted reading times at this position and later, with higher rates of stimulus-matching completions associated with shorter reading times). We do not report these initial regressions here because a reviewer suggested that we also include the condition-wide factors (active/passive structure and Animacy context) in the regressions, and we have followed this suggestion. To foreshadow our results below, we again find reliable effects of completions and also these two additional factors, meaning that the nature of the context (whether an animate or inanimate head noun was promoted, and presence/absence of *bei*) affected reading times beyond the item-specific context effects assessed by the completions. We expect that there are two reasons for these patterns. First, the completion study had fewer observations per cell than the reading data, and so it could be subject to noise if some participants did not read as carefully as the reading time participants, who had a moving-window display and a comprehension question after each sentence. Noise in the completion data from some participants skimming the fragments and providing generic completions would allow the condition-wide Animacy and Structure factors to account for additional variance in reading times. Second, the words in the Animate vs. Inanimate conditions diverge at the head noun and subsequent critical words. At these later words, the Animacy factor, which at that point captures not only the sentence-initial animate/inanimate context but also whether the head noun in fact turned out to be animate or inanimate, could also explain variance of reading time at main clause words, because it is providing additional context that was not seen by participants in the completion task.

The statistical model was built with the log reading time residuals of the corresponding item obtained in the self-paced reading study as the dependent measure. The fixed effects were (1) the percentage of relative clause completions at the Early Gate for the given item, (2) the percentage of relative clause completions at Late Gate that had the same head noun animacy as in the corresponding comprehension item, (3) the type of structure (active vs. passive) and (4) the animacy of the head noun. We ran a separate model for each word starting from the relative pronoun *de*, which strongly disambiguated the stimuli in favor of a relative clause interpretation. Only the Early Gate completions were used as a predictor in the model for the reading times at *de* because the Late Gate was after the *de*. Subjects and items were random effects with random slopes.

No effects were found at reading times for *de*, but we did find effects of these factors at later words, which are shown in Table 4. At the head noun position immediately following the end of the relative clause, Early Gate completions significantly predicted log RT residuals in a negative trend: higher rates of relative clause completions at the Early Gate predicted shorter reading times, and the effect of the Late Gate (which assessed the rates of completions with matching head noun animacy) approached significance at this point (see Table 4 for values). The effects of active/passive structure (the absence/presence of *bei*) and

Table 4
Predicting reading times from completions and stimulus factors.

Factor	Sentence position		
	Head noun	Adverb	Main verb
Early Gate	$\beta = -0.106$, $SE = 0.045$, $t = -2.28$		$\beta = 0.16$, $SE = 0.054$, $t = 2.94$
Late Gate	$\beta = -0.067$, $SE = 0.037$, $t = -1.83$	$\beta = -0.133$, $SE = 0.052$, $t = -2.54$	$\beta = -0.19$, $SE = 0.045$, $t = -4.19$
Structure			$\beta = -0.11$, $SE = 0.028$, $t = -3.94$
Animacy		$\beta = -0.05$, $SE = 0.023$, $t = -2.03$	$\beta = -0.08$, $SE = 0.026$, $t = -3.27$

Note: Early Gate = rate of sentence completions at the Early Gate with relative clause interpretation; Late Gate = rate of sentence completions at the Late Gate with relative clause interpretation with same head animacy as in the full sentence; Structure = Active vs. Passive relative clause structure (absence/presence of passive marker *bei*); Animacy = Sentence-initial context promoting animate or inanimate head noun and (at Head Noun position and later) the animate head noun and associated words in the main clause. Empty cells reflect non-significant results. SE = Standard Error. Negative *t*-values reflect higher completion values associated with lower reading times.

Animacy were not reliable at this word. At the post-head adverb, the Early Gate effects were no longer reliable, but Late Gate continued to be a significant predictor of reading times, as was the effect of Animacy, with inanimate-headed relatives having shorter reading times at this position. However, the effect of active/passive structure, which was strong in the reading time analyses in Experiment 3, did not account for additional variance. At the main verb, Table 4 shows that all four factors were reliable predictors. Here the active/passive Structure factor was reliable for the first time, with reading times shorter in the Passive condition. The other factors had reliable effects in the same direction as at the previous word positions, except that the Early Gate effect is reversed at the main verb, with high rates of relative clause completions associated with higher reading times. With four factors contributing, it is difficult to know how to interpret this result, especially given the consistent effects of Late Gate predictions and that in our initial analyses regressing only the completion data on reading times, the Early Gate completions were not a reliable predictor at this position.

Discussion

The results of the comprehension study and the links to sentence completions argue for comprehenders' sensitivity to fine grained information about the lexico-syntactic properties of relative clauses in Mandarin. Beginning at the head noun—the first position where effects of expectations for head animacy can be tested, the completion data reliably predicted comprehenders' reading times. Both the Early and Late Gate completions were reliable predictors here, but the two dichotomous factors, active/passive structure and Animacy, were not. This result is important in showing that fine-grained information about individual sentences, estimated by completions, reliably predict reading times at the first word for which they can be observed.

Consistent with the results at the head noun, the Late Gate completions continued to be a reliable predictor at the next two words, the adverb and main verb. We interpret the effects of the Animacy factor reading times at these two words as largely reflecting the fact that the stimulus words vary at these positions in the Animate and Inanimate conditions, and only this factor captures that difference.

A second important result is that the reading time data show no evidence for an effect of dependency distance (Gibson, 1998, 2000). Passive relatives have a longer dependency distance than the object relatives and thus should be more difficult than the object relatives, but the passives were read at the same rate or more quickly than the object relatives, and in the analyses including completion data, passives were again associated with shorter reading times at the main verb. Thus we see a clear effect of experience on comprehension, as shown by the relationship between the completions and reading times, but dependency distance makes the wrong predictions for these results.

General discussion

We examined the link between production and comprehension of relative clauses in Mandarin, a language with unique word order and controversial previous comprehension results. We aimed to use Mandarin relative clauses to evaluate two schools of processing accounts, one of which predicts difficulty via inherent properties of the sentence, and one that predicts difficulty via properties of comprehenders' past experiences. The inherent-difficulty accounts explain comprehension difficulty through syntactic complexity, which induces demands on memory maintenance (e.g. Gibson, 1998; Hsiao & Gibson, 2003), while experience-based accounts relate comprehension to learning over language statistics, which Gennari and MacDonald (2008, 2009, MacDonald, 2013) traced to production choices that stem from the nature of grammatical encoding in language production. Our results from a picture description task, corpus analysis, a sentence completion task, and a self-paced reading task show a clear linkage between what speakers produce and what comprehenders expect and interpret in comprehension, modulated by the interaction between structure and the lexico-semantic factor of animacy, instead of an effect of syntactic complexity. These results have implications for theories of both production and comprehension.

Implications for production

Experiment 1 showed clear effects of head noun animacy on speakers' choices of relative clause form, replicating results using a similar task and materials in several typologically diverse languages with head-first relative clauses (Gennari et al., 2012; Montag & MacDonald, 2014, 2015). These results argue for simultaneous planning of the head noun and relative clause despite the fact that

the head noun comes later in the utterance, and thus they argue against extreme forms of incrementality in which the lag between planning and execution is very short (De Smedt, 1996; Kempen & Hoenkamp, 1987; Van Nice & Dietrich, 2003). In our picture description task, the modified head noun had to be identified early in order for the participants to formulate a modifying relative clause, even though the head noun is uttered after the relative clause. The animacy effect of the head noun on structure choice is clear evidence against radical incrementality, as it shows that conceptual properties of a not-yet-produced element influence the sentence structure of a prior clause. Moreover, despite the fact that object and passive relatives in Mandarin have identical word order, and that uttering the very accessible passive marker *bei* can unconditionally lengthen the planning time of later elements, people chose not to do so all the time. The fact that participants produced passives, and some omitted agents, as a function of head noun animacy suggests that both the nouns in relative clause and the head noun were planned in advance. These results favoring a broader scope of planning are consistent with other recent works in languages in which sentence subjects appear late in utterances (Norcliffe, Konopka, Brown, & Levinson, 2015).

The results also support an important role for similarity-based interference in sentence-level production. Interference between semantically similar words (or concepts) is known to interfere with lexical retrieval and increase disfluencies in single-word production (e.g., Glaser & Düngelhoff, 1984; Howard et al., 2006; Oppenheim et al., 2010; Smith & Wheeldon, 2004). At the sentence level, similarity-based interference appears to affect speakers' implicit structure choice in Experiment 1, where more passives, and more agentless passives were produced when both relative clause agent and patient (head noun) were animate (and thus more similar) than when the agent was animate and the head noun was inanimate. Previous accounts of agent omission in Mandarin have suggested that agent omission has simply a discourse origin, where agents are omitted when agent information is given, unimportant or unknown (e.g. Shei, 2014). Discourse factors are of course important in production choices, but our results show that they are not the whole story. In our production task, agents were never mentioned in the prior discourse and were always pictured, and yet they were omitted more when agents acted on animate than inanimate entities. Together, these findings suggest that similarity-based interference has an important role to play in choices of utterance form above and beyond discourse factors, and future research should aim to understand the nature of this interference and its effects. A good starting place would be the manipulation of similarity on a continuum as in Gennari et al. (2012), going beyond the animacy dichotomy here.

The similarity-based interference effects are particularly interesting in that they suggest that speakers are weighing factors from several noun phrases in settling on the form of the utterance (e.g., Stallings & MacDonald, 2011). However, one factor that does not seem to guide production is dependency distance itself, as there is no evidence that Mandarin speakers avoid longer dependencies

(the passive relative) in favor of the shorter dependency object relatives—in fact they favor the passive. MacDonald et al. (in press) have argued that there also is no evidence for dependency distance as a force in production of English relative clauses (cf. Scontras, Badecker, Shank, Lim, & Fedorenko, 2015). Thus while similarity-based interference is hypothesized to stem from the nature of memory representations in the utterance plan, this account is not simply moving existing memory-based comprehension approaches to production, because the classic syntactic complexity and dependency distance effects that are typically invoked in inherent-difficulty (memory limitation) comprehension accounts make the wrong production predictions in Mandarin, where the longer dependency passive is preferred in the present data.

Implications for comprehension–production linkages

Although much work remains to be done, we see our results as most straightforwardly following from the Production–Distribution–Comprehension Account (Gennari & MacDonald, 2009; MacDonald, 2013): production processes favoring accessible animate nouns in prominent grammatical positions and utterance forms that reduce similarity-based interference conspire to yield more passive relatives for animate heads than inanimate heads. Perceivers learn the statistics of these distributions and interpret subsequent input with expectations that this input will conform to past experience. Comprehension difficulty, as measured by reading times, is low when the input matches these expectations and higher when it does not.

A key goal for future research is to better specify the detail over which language users can learn and use the distributional regularities of their language. Our data here showed quite different patterns of passive use across the spoken production task and the corpus analysis of news stories. We also saw that comprehenders track statistics closely in a fine-grained constraint-satisfaction manner, as evidenced by effects of completions on reading times. However, the specificity of this information use is not yet clear. For example, Montag and MacDonald (2015) found that speakers' choices of sentence structure in a spoken task were affected by the different statistics of structure use in written language, which suggests that speakers are influenced to a measureable degree by language usage in other genres. Relatedly, Willits, Amato, and MacDonald (2015) found that comprehenders weighed statistics about events in the world and about lexical co-occurrences to differing degrees depending on task demands. Together, these results suggest that language users have very fine-grained information but also generalize across genres, and they flexibly apply statistical information depending on task demands. Underspecification of these processes is a limitation in accounts positing fine-grained constraint use in part because it is difficult to identify the myriad factors that may influence comprehension and generate strong hypotheses. Computational models of comprehension (e.g. Hsiao & MacDonald, 2013), in which the model learns to weigh information via experience with the input, are an important component to progress.

Steps toward distinguishing inherent-difficulty and experience-based comprehension accounts

The comparison between the object relatives and passive relatives in the present study provides evidence against Dependency Locality Theory, because sentences with longer dependency distance (passive relative clauses) were comprehended more easily than object relative clauses. While Dependency Locality Theory is the memory account most closely associated with Mandarin relative clauses (e.g. Hsiao & Gibson, 2003), the present results are also relevant to other memory-based theories. Cue-based retrieval theories (Gordon et al., 2001; Lewis & Vasishth, 2005; Lewis et al., 2006; Van Dyke & McElree, 2006) are similar in many respects to the Dependency Locality accounts but also posit that similarity-based interference disrupts comprehension. In this account, information from the recent sentence context is retrieved from memory during processing, and the presence of several similar elements in memory disrupts accurate retrieval and processing is slowed. These claims for similarity-based interference in comprehension, together with the effects of similarity-based interference on production in Experiment 1, emphasize the essential challenge of distinguishing inherent-difficulty and experience-based views. As we noted in the introduction, patterns of comprehension difficulty do not by themselves entail a particular theoretical interpretation. For example, long dependencies or inter-item similarities in comprehension materials may be inherently difficult for comprehension processes, because of how memory may inherently be used in comprehension and its decay and interference properties (for discussion, see Van Dyke & Johns, 2012). Alternatively, difficulty might stem from certain stimulus features being rare in comprehenders' prior experience because they are avoided by language producers. Or both could be true—similarity-based interference and/or long dependency distance could be inherently difficult in comprehension, and that difficulty could be exacerbated by the fact that such sentences are rare in comprehenders' experience, as producers avoid them. Clearly the currently available data in the field do not fully discriminate these alternatives, but we sketch some considerations here.

One path to reconciling dependency distance accounts and experience-based accounts is to hypothesize that dependency distance guides production, such that producers avoid syntactically complex utterances and longer dependency distances in favor or simple, short-distance alternatives. Scontras et al. (2015) suggested that longer dependency sentences may yield production difficulty, but MacDonald et al. (in press) argued that other factors (including producers avoiding similarity-based interference) conspire to discourage certain sentence forms that also happen to have long dependencies, and that there is not an independent effect of dependency distance in production in English. For Mandarin, the production results from Experiment 1 also argue against any effect of dependency distance in production, because speakers preferred the longer-distance passive structures.

Alternatively, comprehension difficulty for sentences with semantic similarity may stem from these sentences being rare, because they are avoided by producers, instead of, or in addition to effects of similarity within the comprehension system. Future research aimed at distinguishing alternative comprehension accounts should therefore include corpus analyses to provide broader data about the extent to which sentences that could engender similarity-based interference are truly rarer than sentences in which the relevant words are less similar. Hsiao et al.'s (2014) corpus analysis was a start in this direction but was relatively small. Additional comprehension studies attempting to pin down the nature of interference in comprehension (e.g. Kush, Johns, & Van Dyke, 2015) could also be useful in identifying whether there is an inherent difficulty component to similarity in comprehension in addition to the experience-based difficulty that likely exists. In terms of our own self-paced reading results in Experiment 3, the difficulty of animate-headed active object relative clauses could be interpreted as an effect of similarity-based retrieval interference in comprehension, as both the head noun and relative clause noun are animate and similar. However, this interpretation is not consistent with the passive reading times in the same study, where both animate-headed passives, with semantically similar words, were no harder than the inanimate passives. Thus similarity-based interference in comprehension does not seem to add an account of these data. Our own bias is therefore to lodge the effects of similarity-based interference in production and view the comprehension difficulty results as largely experience-dependent (see MacDonald, 2013, for why production could be the primary locus of difficulty). However, the space of comprehension–production linkages is not fully explored to date.

Other comprehension considerations: ambiguity and surprisal

A number of researchers have noted that inherent difficulty accounts are aimed at explaining processing difficulty of unambiguous sentences, via processing costs associated with syntactic complexity or other factors (Gennari & MacDonald, 2008; Levy, Fedorenko, & Gibson, 2013; MacDonald, 2013). Ambiguity significantly clouds interpretation of comprehension results, because ambiguity itself causes processing difficulty. The abundance of ambiguity in Mandarin relative clauses makes it difficult to ascribe comprehension difficulty to one factor or the other. However, we question the notion that there are any truly unambiguous sentences (MacDonald, 2013): as researchers working within surprisal/entropy accounts have noted, there is always some degree of uncertainty about the upcoming input (Hale, 2001; Levy, 2008). Whereas Levy et al. (2013, p. 462) suggest that "theories of syntactic complexity…may be able to subsume theories of ambiguity resolution" in accounts of comprehension difficulty, we favor the opposite approach—if all linguistic input contains ambiguity, in the sense of uncertainty about upcoming input and possibly other senses, then it may be possible to subsume accounts of syntactic complexity, dependency distance, etc. under (experience-dependent) processes of expectation and ambiguity resolution.

To address the ambiguity-dependent confound for Mandarin relative clauses and evaluate inherent difficulty

accounts, Jäger et al. (2015) embedded Mandarin relative clauses in sentence contexts that greatly reduced syntactic ambiguities. They found evidence that processing difficulty was linked to the amount of experience comprehenders had with various relative clause types, and not with their syntactic complexity. These results are in many respects compatible with our own approach, but Jäger et al.'s perspective differs from ours in one crucial respect that is instructive in distinguishing our own approach from other experience-based approaches, particularly Surprisal (Hale, 2001; Levy, 2008). Surprisal is the degree to which a word is unexpected in a sentence, but the relevant grain(s) over which expectations are developed can take many forms. Jäger et al. view experience in terms of what Demberg and Keller (2008) call "unlexicalized surprisal", calculating expectations over grammatical categories rather than individual words, so that expectations are independent of lexical content. By contrast, we have emphasized a much finer grain of experience-based expectations, in that we have shown how head noun animacy, and indeed the lexical and discourse properties of individual sentences, affect comprehenders' expectations (as measured by completions) and reading times. Consistent with this view, Gennari and MacDonald (2008) regressed different aspects of their English completion data on comprehenders' reading times to investigate the grains at which expectations appeared to be forming. They found that expectations for a particular thematic role (which incorporates syntactic and lexical properties of nouns and verbs) was a good predictor of reading times, as were predictions for the exact word in some regions. By contrast, lexically-independent sentence structure (dependency distance) did not account for additional variance.

A second way in which our account differs from Jäger et al.'s (2015) is in the calculation of the relevant experiences that generate expectations. Hsiao and MacDonald (2013) emphasized not only the importance of lexical information such as animacy and animacy–structure co-occurrences but also the "neighborhoods" and sub-regularities in sentence structures affect processing difficulty in Mandarin relative clauses (see also Fitz, Chang, & Christiansen, 2011; MacDonald & Christiansen, 2002). On this view, the difficulty of a structure, such as a relative clause, is dependent not only on the frequency of that structure and its lexico-syntactic co-occurrences (such as the pairing of a structure and nouns of a certain animacy), but it is also dependent on the extent to which that structure shares properties with other structures in the language. Computational models, such as the simple recurrent networks (SRNs) used by Fitz et al., MacDonald and Christainsen, and Hsiao and MacDonald, are an important component in investigating this hypothesis, because the models, trained on a range of sentence types, learn the neighborhood structures and other regularities. As a result, processing difficulty for a given sentence type, such as an object relative clause, depends on more than simply the frequency of that sentence type in the input. For example, some more frequent structures yield higher error rates in Hsiao and MacDonald's (2013) model (and longer reading times for human comprehenders) than some rarer structures (e.g., Mandarin subject-modifying subject relative clauses are more frequent than object relatives in the corpus but still harder to process). Whereas previous researchers might have used evidence of this frequency-difficulty asymmetry as evidence against experience accounts, our view is that it is evidence only against overly simple versions of those approaches, and it is instead evidence for neighborhood effects in sentence comprehension. Indeed, in his comparison of surprisal and SRNs in accounting for comprehension difficulty, Frank (2009) argued that though unlexicalized surprisal provides a better language model (that is, it is a more veridical calculation of the frequency of structures in a corpus), the SRNs provide a better account for patterns of human comprehension difficulty: both humans and the models are affected by complex neighborhoods and sub-regularities that are not captured by the raw frequency of a certain sentence type. Because neighborhoods also affect production as well as comprehension (Fitz et al., 2011), we expect that this approach will be important in continuing to explore comprehension–production links in relative clauses and other facets of language use.

Appendix A. TGrep2 search patterns for passive relative clauses

Full passive relatives:
/^NP/<<(CP   <<(/^IP/<(/-SBJ/</^-NONE-/)<<(/^VP/<<L B<<VV<<(/-OBJ/</^-NONE-/))$..DEC))
Bare passive relatives:
/^NP/<<(CP   <<(/^IP/<(/-SBJ/</^-NONE-/)<<(/^VP/<<S B<<VV<<(/-OBJ/</^-NONE-/))$..DEC))

Appendix B. Experimental items for Exp. 3

Each sentence begins with a context phrase, followed by an object relative clause or a passive relative clause that modifies the main clause subject. (Only the object relative clauses are shown here. Passive relatives only require the addition of bei at the beginning of the relative clause.)

	Animate	Inanimate
1	为了帮媒体讨回一笔公道, 电影明星打了的记者愤怒地告上法庭。	虽然有好几百斤重, 电影明星打了的沙包夸张地破了洞。
	To demand justice for the media, the reporter that the movie star beat up angrily raised a legal complaint.	Despite weighing hundreds of kilograms, the sandbag that the movie star beat was exaggeratedly broken with a hole.

Appendix B (*continued*)

	Animate	Inanimate
2	为了跻身进入皇室家族, 王子注视的平民女孩们卖力地表现。 In order to marry into the royal family, the civilian girls that the prince was gazing at really put out their energy to show their best.	经过了浩大的工程, 王子注视的城堡气派地座落于花园的中央。 After a huge scale construction, the castle that the prince was gazing at was situated in an imposing manner at the center of a garden.
3	受不了身體的疼痛, 营救队搬运的伤者痛苦地哀嚎。 Unable to put up with the pain, the injured people that the rescue team was carrying cried out in pain.	因为一直没有换旧更新, 营救队搬运的担架老旧地发出嘎嘎声响。 Not having been replaced for new ones, the stretcher that the rescue team was carrying made noisy sounds.
4	虽然伤得不重, 邮差撞到的路人嚣张地索赔。 Even though not hurt badly, the pedestrian that the postman crashed into aggressively asked for compensation.	因为停放的位置超出停车格, 邮差撞到的小客车严重地损坏。 Parked outside of the parking space, the sedan that the postman crashed into was severely damaged.
5	因为站得太靠近球场, 球员击中的观众疼痛地大叫出声。 Standing too close to the field, the spectator that the play hit shouted out in pain.	飞离场外后, 球员击中的棒球意外地砸中观众。 After flying out of the field, the baseball that the player hit accidentally hit a spectator.
6	因为好不容易找到工作, 顾客抱怨的店员卑微地请求原谅。 Having landing this job after a long search, the clerk that the customer complained about humbly asked for forgiveness.	虽然经过严格的管控, 顾客抱怨的商品意外地出了瑕疵。 Even with strict quality control, the product that customers complained about was surprisingly flawed.
7	躺入了箱子之后, 魔术师触摸的女助手神奇地凭空消失。 After lying down in the box, the female assistant that the magician touched magically disappeared.	纵使体积巨大, 魔术师触摸的箱子如常地凭空消失。 Even with its huge volume, the box that the magician touched still disappeared just as usual.
8	为了不让法官起疑, 受害者怒瞪着的嫌疑犯流畅地提供不在场证明. To not make the judge suspicious, the suspect that the plaintiff stared at offered his alibi fluently.	因为清楚地印有犯人的指纹, 受害者怒瞪着的证物血淋淋地证实犯人的罪刑。 Marked with the suspect's finger prints, the exhibit that the plaintiff stared at served as smoking gun evidence for the crime of the convict.
9	因为不习惯被别人注视, 画家临摹的小女孩害羞地动来动去。 Not used to being stared at, the girl that the painter was painting bashfully moved around.	整齐地排列在桌子上, 画家临摹的静物完美地提供最好的视角。 Lining up tidily on the table, the objects that the painter is drawing provides the best angle.
10	为了要保持全勤, 室友摇晃的女孩痛苦地醒来。 To keep full time attendance, the girl that the roommate shook woke up unwillingly.	原本就因为老旧而不稳, 室友摇晃的椅子吵闹地发出吱吱声。 Old and unstable, the chair that the roommate is rocking noisily makes sounds.
11	因为不愿意引起轩然大波, 媒体批评的歌手低调地不做任何回应。 Unwilling to cause an uproar, the singer that the media criticized made no comments in a low-key way.	虽然剧情沉闷了无新意, 媒体批评的电影意外地获得大奖。 Although the plot was dull and unoriginal, the movie that the media criticized surprisingly won a major award.
12	因为赶时间, 交警引导的驾驶员嚣张地闯红灯。 Being in a rush, the driver that the traffic police were guiding daringly ran through the red light.	一辆接着一辆, 交警引导的车辆顺利地通过十字路口。 One car after another, the vehicles that the traffic police were guiding smoothly crossed the intersection.
13	使出了浑身解数, 凶手压倒的人质敏捷地挣脱出来。 After trying everything he could, the hostage that the criminal pressed down nimbly escaped.	上面放着许多货品, 凶手压倒的货架沉重地落在警察身上。 With many goods on top, the shelf that the criminal pushed down fell heavily on the police.
14	因为长相是全村最美, 祭司献上的女孩牺牲地成为结束灾难的祭品。 Being the prettiest in the village, the girl that the priest offered became the offering in hope to end the disaster.	洗干净之后, 祭司献上的祭品神圣地放置在祭坛上。 After being washed clean, the offering that the priest presented was placed sacredly on the altar.
15	因为民族优越感很强, 外交官访问的邦交国官员傲慢地回答问题。 Being very ethnocentric, the officials from the country with diplomatic relation that the diplomat visited answered questions arrogantly.	虽然不如城市的繁华, 外交官访问的乡村完整地体现温暖的人情味。 Though not a bustling city, the village that the diplomat visited completely showed heartwarming hospitality.

Appendix B (*continued*)

	Animate	Inanimate
16	遭抢之后, 抢匪推倒的老太太慢慢地爬起身。 After being robbed, the old lady that the bandit pushed down slowly got up.	杂乱地散落一地, 抢匪推倒的障碍物成功地绊倒追赶者。 Scattered around on the floor, the obstacle that the bandit pushed down successfully tripped the people chasing after.
17	觉得自己很幸运, 主持人抽中的中奖者雀跃地上台领奖。 Feeling lucky, the lottery winner that the host drew out went on the stage excitedly to receive the prize.	在所有的奖项中, 主持人抽中的头奖抢眼地获得最大的尖叫声。 Among all the items, the first prize that the host drew out eye-catchingly stirred the loudest cheer.
18	因为体力不支, 枪手射中的敌人虚弱地喘息着。 Out of strength and energy, the enemy that the gunman shot was weakly breathing.	尽管红心很小, 枪手射中的枪靶迅速地倒下。 With a very small red center, the target that the shooter shot at quickly went down.
19	因为太多赌债还没讨回来, 男人躲着的债主凶狠地找上门来。 With so much debt not repaid, the creditor that the man had been avoiding viciously came to the door and demand repayment.	笔直地飞出后, 男人躲着的飞箭稳稳地插中他的心脏。 Flying straight out, the arrow that the man was avoiding went right into his heart.
20	接受了一小时的疗程后, 美容师按摩的上班族渐渐地释放平日累积的压力。 After receiving one hour of treatment, the salary man that the beautician massaged gradually released stress.	原本红肿酸痛不止, 美容师按摩的部位迅速地得到释放。 Originally being sore and swollen, the body part that the beautician massaged soon was relieved.
21	因为受过训练, 摄影师拍摄的模特专业地摆着姿势。 Having received training, the model that the photographer was shooting professionally posed for the shots.	现场百花齐放, 摄影师拍摄的风景美丽地展露生机。 With flowers blooming, the natural scene that the photographer was shooting gracefully exhibited vitality.
22	跌倒了之后, 高中生扶着的老先生慢慢地爬起身。 After tumbling off, the old man that the high school student helped support slowly got up.	因为年久失修, 大学生扶着的楼梯扶手斑驳地挂满铁锈。 Worn down by years of non-repair, the stair railing that the college student supported himself on was full of rust scales.
23	因为事先预习过, 英文老师点到的男同学拿手地回答着问题。 Having previewed the materials, the male student that the English teacher called on answered the question very well.	因为充斥着难以理解的文字, 英文老师点到的课文可想而知地难倒了所有人。 Full of incomprehensible words, the text that the English teacher called on obviously baffled everybody.
24	中枪了之后, 警察追捕的逃犯虚弱地喘息着。 After being shot, the criminal that the police chased after was panting weakly.	因为发出的噪音超过标准限度, 警察追捕的摩托车飞快地逃逸。 Emitting noises that were over the limit, the motorcycle that the police chased after got away at a lightning speed.
25	因为不需要自己走路, 小男孩背着的小女孩开心地唱着歌。 Not having to walk by herself, the little girl that the little boy carried on the back was happily singing.	因为装了许多的书本, 小男孩背着的书包沉重地压着他的身子。 Packed with a lot of books, the bag that the little boy carried on the back heavily burdened his body.
26	因为突如其来的疼痛, 男孩踢到的小妹妹嚎啕地大哭起来。 Because of the sudden pain, the young girl that the boy kicked started bawling.	在强劲的力道下, 男孩踢到的足球快速地滚进球门。 Under the strong force, the soccer ball that the boy kicked rapidly rolled in to the goal.
27	因为来不及逃, 军队袭击的人民悲惨地无家可归。 Too late to escape, the civilians that the troops attacked sadly became homeless.	因为地处偏远, 军队袭击的村庄悄悄地变成一座鬼城。 Located at the remote place, the village that the troops attacked quietly became a ghost town.
28	泄密之后, 警官查捕的间谍已经束手就擒。 After leaking top secrets, the spy that the police officer hunted down was already arrested.	遭窃了许久后, 警官查捕的赃车准确地归还原主。 Having been stolen a long time ago, the car that the police officer hunted down was handed back to the owner precisely.

Appendix C. Statistical results of mixed-effects models of word-by-word reading times in Exp. 3

Significant fixed effects of interest (non-spillover effects) are marked with *$^{(p<.05)}$.

Fixed effects	Coefficient	SE	t	p
Word 1 (bei)				
Animacy = Inanimate	0.055	0.023	2.36	.02
Spillover n-1	−0.072	0.033	−2.14	.03
Spillover n-2	0.052	0.039	1.34	.18
Spillover n-3	0.126	0.043	2.97	.003
Word 2 (RC-internal noun)				
Animacy = Inanimate	0.029	0.023	1.22	.22
Structure = Passive	−0.033	0.023	−1.44	.15
Interaction = Inanimate & Passive	−0.022	0.033	−0.66	.51
Spillover n-1	0.083	0.024	3.51	.0005
Spillover n-2	0.060	0.025	2.45	.015
Spillover n-3	0.153	0.027	5.63	<.0001
Word 3 (RC-internal verb)				
Animacy = Inanimate	0.005	0.022	0.21	.83
Structure = Passive	0.006	0.022	0.29	.77
Interaction = Inanimate & Passive	0.012	0.031	0.37	.71
Spillover n-1	0.059	0.022	2.63	.009
Spillover n-2	0.096	0.022	4.34	<.0001
Spillover n-3	0.021	0.022	0.97	.33
Word 4 (DE)				
Animacy = Inanimate	−0.004	0.019	−0.20	.84
Structure = Passive	−0.013	0.019	−0.72	.47
Interaction = Inanimate & Passive	0.007	0.026	0.25	.80
Spillover n-1	0.040	0.021	1.92	.05
Spillover n-2	0.054	0.019	2.84	.005
Spillover n-3	0.035	0.019	1.84	.07
Word 5 (head noun)				
Animacy = Inanimate	−0.019	0.021	−0.91	.36
Structure = Passive	−0.034	0.021	−1.59	.11
Interaction = Inanimate & Passive	0.018	0.030	0.60	.55
Spillover n-1	0.269	0.031	8.77	<.0001
Spillover n-2	0.039	0.023	1.64	.10
Spillover n-3	0.095	0.022	4.36	<.0001
Word 6 (post-head adverb)				
Animacy = Inanimate	−0.068	0.022	−3.16	.002
Structure = Passive	−0.060	0.022	−2.77	.006
Interaction = Inanimate & Passive	0.077	0.031	2.52	.012
Spillover n-1	0.062	0.022	2.77	.006
Spillover n-2	0.155	0.032	4.84	<.0001
Spillover n-3	0.064	0.024	2.71	.007
Word 7 (matrix verb)				
Animacy = Inanimate	−0.110	0.023	−4.75	<.0001
Structure = Passive	−0.114	0.023	−4.90	<.0001
Interaction = Inanimate & Passive	0.097	0.033	2.97	.003
Spillover n-1	0.099	0.025	3.87	.0001
Spillover n-2	0.026	0.024	1.11	.268
Spillover n-3	0.012	0.034	0.34	.731

References

Bock, K. (1987). An effect of the accessibility of word forms on sentence structures. *Journal of Memory and Language, 26*, 119–137.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*, 47–67.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chen, B., Ning, A., Bi, H., & Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica, 129*, 61–65.

De Smedt, K. (1996). Computational models of incremental grammatical encoding. In T. Dijkstra, K. de Smedt, T. Dijkstra, & K. de Smedt (Eds.), *Computational psycholinguistics: AI and connectionist models of human language processing* (pp. 279–307). Philadelphia, PA, US: Taylor & Francis.

Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition, 109*, 193–210.

Drieman, G. H. J. (1962). Differences between written and spoken language. *Acta Psychologia, 20*, 36–57. 78-100.

Fedorenko, E., Woodbury, R., & Gibson, E. (2013). Direct evidence of memory retrieval as a source of difficulty in non-local dependencies in language. *Cognitive Science, 37*, 378–394.

Ferreira, V. S., & Firato, C. E. (2002). Proactive interference effects on sentence production. *Psychonomic Bulletin & Review, 9*, 795–800.

Fitz, H., Chang, F., & Christiansen, M. H. (2011). A connectionist account of the acquisition and processing of relative clauses. In E. Kidd (Ed.). *The acquisition of relative clauses* (Vol. 8, pp. 39–60). Amsterdam: John Benjamins.

Frank, S. L. (2009). Surprisal-based comparison between a symbolic and a connectionist model of sentence processing. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the Cognitive Science Society* (pp. 1139–1144). Austin, TX: Cognitive Science Society.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language, 58*, 161–187. http://dx.doi.org/10.1016/j.jml.2007.07.004.

Gennari, S. P., & MacDonald, M. C. (2009). Linking production and comprehension processes: The case of relative clauses. *Cognition, 111*, 1–23. http://dx.doi.org/10.1016/j.cognition.2008.12.006.

Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology, 65*, 141–176. http://dx.doi.org/10.1016/j.cogpsych.2012.03.002.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*, 1–76.

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

Gibson, E., & Wu, H.-H. I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes, 28*, 125–155.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 640.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1411.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language, 51*, 97–114.

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentenial complexity. *Cognitive Science, 29*, 261–290.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the 2nd meeting of the North American chapter of the Association for Computational Linguistics* (pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition, 100*, 464–482. http://dx.doi.org/10.1016/j.cognition.2005.02.006.

Hsiao, Y., Gao, Y., & MacDonald, M. C. (2014). Agent-patient similarity affects sentence structure in language production: Evidence from subject omissions in Mandarin. *Frontiers in Psychology, 5*, 1015. http://dx.doi.org/10.3389/fpsyg.2014.01015.

Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition, 90*, 3–27.

Hsiao, Y., & MacDonald, M. C. (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Frontiers in Psychology, 4*(767), 2013. http://dx.doi.org/10.3389/fpsyg.00767.

Hundt, M. (2004). Animacy, agentivity, and the spread of the progressive in modern English. *English Language and Linguistics, 8*, 47–69.

Jaeger, T. F. (2005). Optional that indicates production difficulty: Evidence from disfluencies. In *Disfluency in spontaneous speech* (pp. 103–108).

Jaeger, T. F., & Wasow, T. (2005). Production complexity driven variation: The case of relativizer distribution in non-subject-extracted relative clauses. In *The 18th annual CUNY sentence processing conference*, March 31st–April 2nd.

Jäger, L., Chen, Z., Li, Q., Lin, C.-J. C., & Vasishth, S. (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language, 79–80*, 97–120. http://dx.doi.org/10.1016/j.jml.2014.10.005.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*, 201–258.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*, 580–602.

Kush, D., Johns, C., & Van Dyke, J. (2015). Identifying the role of phonology in sentence-level reading difficulty. *Journal of Memory and Language, 79*, 18–29.

Levelt, W. J. (1989). *Speaking: From intention to articulation. ACL. MIT press series in natural-language processing*. Cambridge, Massachusetts: MIT Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106*, 1126–1177.

Levy, R., Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language, 69*, 461–496. http://dx.doi.org/10.1016/j.jml.2012.10.005.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*, 375–419.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences, 10*, 447–454.

Lin, C.-J. C. (2011). Chinese and English relative clauses: Processing constraints and typological consequences. In Z. Jing-Schmidt (Ed.). *Proceedings of the 23rd North American conference on Chinese linguistics (NACCL-23)* (Vol. 1, pp. 191–199). Eugene: University of Oregon.

Lin, C.-J. C., & Bever, T. G. (2006). Subject preference in the processing of relative clauses in Chinese. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th west coast conference on formal linguistics (WCCFL 25)* (pp. 254–260). Somerville, MA: Cascadilla Press.

Lin, Y., & Garnsey, S. M. (2011). Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin. In H. Yamashita, Y. Hirose, & J. L. Packard (Eds.), *Processing and producing head-final structures* (pp. 241–276). New York, NY: Springer.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00226.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54.

MacDonald, M. C., Montag, J. L., & Gennari, S. P. (2015). Dependency distance does not predict production difficulty: Reply to Scontras et al. *Cognitive Science* (in press).

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676–703.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*, 283–312.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language, 35*, 477–496. http://dx.doi.org/10.1006/jmla.1996.0026.

Montag, J. L., & MacDonald, M. C. (2015). Text exposure predicts spoken production of complex sentences in eight and twelve year old children and adults. *Journal of Experimental Psychology: General, 144*, 447–468.

Montag, J. L., & MacDonald, M. C. (2014). Effects of text exposure on spoken sentence production. In P. Bello, M. Guarini, M. McShane, & B.

Scassellati (Eds.), *Proceedings of the 36th annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience, 30*, 1187–1208.

O'Grady, W. (2011). Relative clauses: Processing and acquisition. In *The acquisition of relative clauses: Processing, typology and function* (pp. 13–38).

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2007). Cumulative semantic interference as learning. *Brain and Language, 103*, 175–176.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114*, 227–252.

Pu, M.-M. (2007). The distribution of relative clauses in Chinese discourse. *Discourse Processes, 43*, 25–53.

Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In *Proceedings of the 25th annual meeting of the Cognitive Science Society*.

Rohde, D. (2001). *Linger*. <http://tedlab.mit.edu/~dr/Linger/>.

Rohde, D. (2005). *TGrep2 manual*. <http://tedlab.mit.edu/~dr/Tgrep2/tgrep2.pdf>.

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

Schallert, D. L., Kleiman, G. M., & Rubin, A. D. (1977). Analyses of difference between written and oral language. Center for the Study of Reading, Technical Report No. 29.

Scontras, G., Badecker, W., Shank, L., Lim, E., & Fedorenko, E. (2015). Syntactic complexity effects in sentence production. *Cognitive Science, 39*, 559–583. http://dx.doi.org/10.1111/cogs.12168.

Shei, C. (2014). *Understanding the Chinese language: A comprehensive linguistic introduction*. New York, NY: Routledge.

Smith, M., & Wheeldon, L. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 675–686.

Stallings, L. M., & MacDonald, M. C. (2011). It's not just the "heavy NP": Relative phrase length modulates the production of heavy-NP shift. *Journal of Psycholinguistic Research, 40*, 177–187.

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition, 114*, 447–454.

Su, Y.-C., Lee, S., & Chung, Y. (2007). Asyntactic thematic role assignment by Mandarin aphasics: A test of the Trace-Deletion Hypothesis and the Double Dependency Hypothesis. *Brain and Language, 101*, 1–18.

Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences in word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language, 65*, 318–330.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.). *Handbook of perception and cognition: Speech, language and communication* (Vol. 11, pp. 217–262). San Diego, CA: Academic Press.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*, 285–318.

Van Dyke, J. A., & Johns, C. L. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass, 6*, 193–211.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence processing. *Journal of Memory and Language, 55*, 157–166.

Van Nice, K. Y., & Dietrich, R. (2003). Animacy effects in language production: From mental model to formulator. In H. Härtl & H. Tappe (Eds.), *Mediating between concepts and grammar* (pp. 101–117). Berlin: Mouton de Gruyter.

Vasishth, S., Chen, Z., Li, Q., & Guo, G. (2013). Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PloS One, 8*, e77006.

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

Willits, J. A., Amato, M. S., & MacDonald, M. C. (2015). Language knowledge and event knowledge in language use. *Cognitive Psychology, 78*, 1–27.

Wu, F. (2009). *Factors affecting relative clause processing in Mandarin*. Unpublished doctoral dissertation, University of Southern California, Los Angeles, CA, USA.

Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes, 27*, 1489–1524.

Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in Contrast, 6*, 109–149.

Xue, N., Jiang, Z., Zhong, X., Palmer, M., Xia, F., Chiou, F.-D., & Chang, M. (2010). *Chinese Treebank 7.0*. Philadelphia, PA: Linguistic Data Consortium.

Yang, C. L., Perfetti, C. A., & Liu, Y. (2010). Sentence integration processes: An ERP study of Chinese sentence comprehension with relative clauses. *Brain & Language, 112*, 85–100.