Comprehension of *Who* and *Which*-NP questions: Which Account do the Data Support?

This study investigated the comprehension of various types of *Wh*-questions in unimpaired adults and those with Broca's aphasia. Using an eye tracking-while listening method, we examined four specific hypotheses. Our initial results for our healthy controls revealed offline support for a Word Order Hypothesis – object-extracted *Who* and *Which*-NP questions took longer to resolve than subject-extracted versions. Our results using eye gaze data, however, revealed support for a Retrieval Hypothesis – *Which*-NP questions that contain more specific information yielded fewer looks to the correct referent than their *Who*-question counterparts. These patterns set the stage for our ongoing aphasia study.

This study examined comprehension of *Wh*-questions in neurologically healthy adult participants and individuals with aphasia. *Wh*-questions are an interesting structure to investigate for several reasons. Consider the following table of examples:

Table 1: Example materials for Wh-questions

Discourse: Yesterday afternoon two mailmen and a fireman had a fight.	
Subject-Who	Who pushed the fireman yesterday afternoon?
Object-Who	Who did the fireman push yesterday afternoon?
Subject-Which	Which mailman pushed the fireman yesterday afternoon?
Object-Which	Which mailman did the fireman push yesterday afternoon?

Table 1 shows two types of questions, *Who* and *Which*-NP, and within each question type there are subject-extraction and object-extraction versions. There is evidence that *Which*-NP questions are more difficult to process for neurologically healthy participants than *Who*-question counterparts because *Which*-NP constructions must refer to an individual from a set of individuals in the discourse, while *Who* questions are not required to do so. We call this the "Discourse Hypothesis" (see Donkers & Stowe, 2006; Shapiro, 2000). Relatedly, there is evidence for the opposite effect (see Hofmeister, 2007) because *Which*-NP constructions, relative to their *Who*-question counterparts, contain specific information that speeds memory retrieval; we call this the "Retrieval Hypothesis." There is also evidence that across the two question types, object-extracted questions are more difficult to understand than subject-extracted questions because the former are in non-canonical word order. We call this the "Word Order Hypothesis."

In aphasia research, similar discrepancies across studies have been observed. For example, Hickok and Avrutin (1996) found that subject *Which* questions were comprehended significantly better than object *Which* questions, yet there was no difference found between subject and object *Who* questions. More recently in the linguistic and child language literatures, a distinction between *Who* and *Which*-NP questions (among other types of constructions) has been made in terms of "Interveners". Briefly here, an intervener is an NP that has similar properties to other NPs in the sentence, and thus interferes with the assignment of thematic roles. Only object *Which*-questions fall under this hypothesis and thus these are suggested to be more difficult to process (Friedmann & Novogrodsky, 2011). Similar predictions can be cast for aphasia, which we call the "Intervener Hypothesis."

Thus, the goal of our investigation was to clear up the empirical inconsistencies found in the literature and to test these various hypotheses using a unique method, which we describe below. Importantly as well, *Wh*-questions are fundamental to functional communication, and indeed have been the focus of some treatment programs in aphasia (e.g., Shapiro & Thompson, 2006; Thompson & Shapiro, 2007).

Method

Participants. Twenty-two healthy control participants and seven adults with agrammatic aphasia participated in the study. All individuals with aphasia had experienced a single, unilateral left hemisphere stroke. Diagnosis of agrammatism was based on the BDAE (version 3; Goodglass, Kaplan, & Barresi, 2000), and the WAB (Kertesz, 2006) as well as chance performance on comprehension of non-canonical (object-relative and passive) sentences from the SOAP Test (Love & Oster, 2002).

Materials and Design. 10 examples of four types of *Wh*-question structures, with felicitous discourse, were developed (see Table 1), yielding a 2X2 design, with *Wh*-question type (*Who* and *Which*-NP) and within those, extraction type (subject, object). Each sentence was coupled with a picture containing three figures interacting with one another (see Figure 1).

Figure 1: Example for "…two mailmen and a fireman got into a fight.
Which mailman/Who did the mailman push?"

Four counterbalanced lists were developed such that any given participant only saw each picture once in each list; participants completed four sessions to receive every stimulus. Filler sentences and accompanying pictures were developed that required a middle referent (e.g., the fireman) response.

Procedure. We used an eye tracking-while-listening method. In the standard version of this method, called "the visual world" (see Sussman & Sedivy, 2003), participants are presented with sentences and look at a rectangle array containing nine squares, with a fixation point in the middle and pictures of the nouns mentioned in the sentence as well as foils that are irrelevant to the sentence placed in four of the squares. As shown in Figure 1, however, we used action pictures that better reflect the dynamics of the sentence and thus gazes to the subject and object referents should more accurately measure comprehension.

Participants sat in front of the eye-tracker while viewing a picture (see Figure 1) and listening to the discourse and subsequent question. The participants were instructed to answer the question using a button box with three buttons labeled "Left," "Right," and "Middle", corresponding to the three figures in each picture, and response times and accuracy were recorded. For the eye-tracking data, gazes to each of the three regions were recorded. The data of interest consisted of the ratio of percentage of time spent gazing at the correct region to the incorrect region in each condition across the time course of the question; future analyses will examine time-locked gazes. Here we only discuss the normal control data (N=22) as we continue to analyze the data from our participants with aphasia.

Results and Discussion

For the normal controls, there were ceiling effects for accuracy; all participants had 90% accuracy or greater. Repeated measures ANOVA and paired comparisons conducted on the RTs from trials with correct responses revealed that RTs for the object *Which* condition were significantly slower than those for the subject *Which* condition, t (21) = -2.81, p = .01, and the RTs for the object *Who* condition were significantly slower than in the subject *Who* condition, t (21) = -2.302, p = .032. These initial data support the Word Order Hypothesis; object-extracted constructions took longer to process than subject-extracted, regardless of Question type.

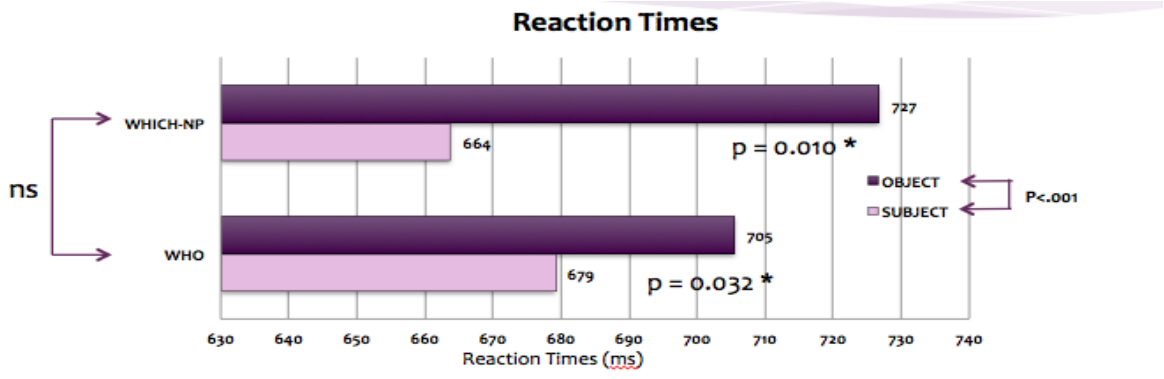

Figure 2. RT Results across Question and Extraction Type

Gaze Data

First, protected paired t-tests were conducted comparing the percentage of gazes to the correct referent to the percentage of gazes to the incorrect referent in each of the four conditions. These were all significant at the $p < .001$ level, serving as a 'reality check' on our method. Next, the ratio of the percentage of looks to the correct referent to the percentage of looks to the incorrect referent was calculated and was used as the dependent measure. A repeated measure ANOVA and paired comparisons revealed that the percentage ratio was significantly higher in the object *Who* condition (4.51) compared to the subject *Who* condition (2.63), $t(21) = 2.392$, $p = 0.026$. There was no significant difference between the object *Which* condition (3.38) compared to the subject *Which* condition (2.30).

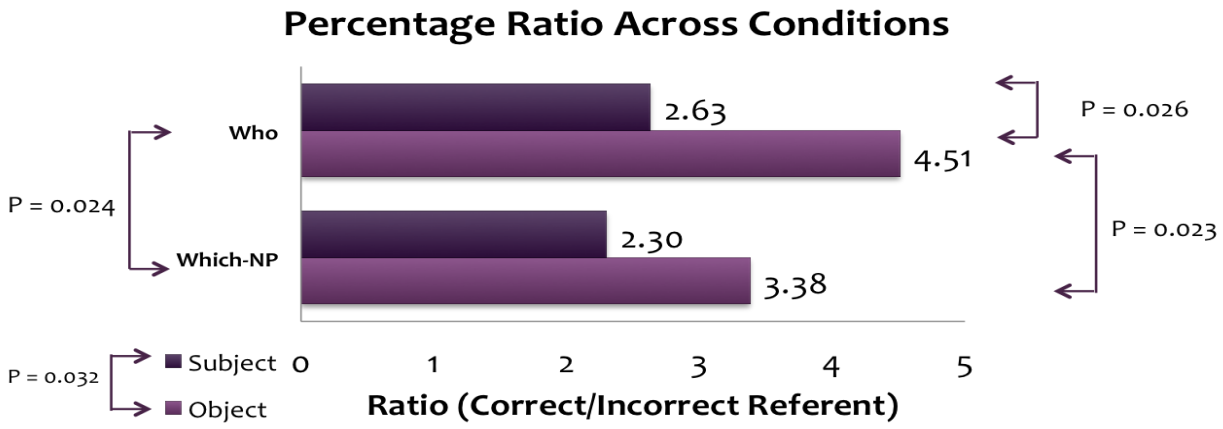

Figure 3. Ratio of percentage of gazes to the correct referent relative to the incorrect referent

These gaze data support the Retrieval Hypothesis. Recall that Hofmeister (2007) found that *Which* NP-questions had a significant processing advantage over *Who* questions, and attributed this distinction to the fact that *Which*-NP questions contain more information than *Who* questions. For example, in the question "Which fireman hit the mailman?" the phrase "*Which fireman…"* is associated with implied features from a person's real-world knowledge of firemen,

whereas *Who* questions do not contain implied features and thus are more difficult to process. Further discussion of our results will center on the distinction found between or offline RT data and our online gaze data. These data serve as background to our aphasia data, which preliminarily show distinct patterns from the control participants.

References

Donkers, J. & Stowe, L. (2006). Wh-questions and the nature of D-linking: A processing perspective. International Conference on Linguistic Evidence, University of Tübingen, Germany.

Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand: The comprehension of Wh questions in three types of SLI. *Lingua, 121,* 367-382.

Hickok, G., & Avrutin, S. (1996). Comprehension of Wh-questions in two Broca's aphasics. *Brain and Language, 52,* 314-327.

Hofmeister, P. (2007). Retrievability and Gradience in Filler-Gap Dependencies. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *43*(1), 109-123.

Shapiro, L.P., & Thompson, C.K. (2006). Training language deficits in Broca's aphasia. In Grodzinsky, Y., and Amunts, K. (editors), *Broca's Region.* Oxford: Oxford University Press.

Shapiro, L.P. (2000). The processing of long-distance dependencies in normal listeners: Evidence for form-driven activation. In R. Bastiannse & Y. Grodzinsky (Eds.), *Grammatical Disorders in Aphasia: A Neurolinguistic Perspective*. London: Whurr Publishers.

Shapiro, L.P. & Thompson, C.K. (2006). Training language deficits in Broca's aphasia. In Y. Grodzinsky and K. Amunts (Eds.), *Broca's Region.* Oxford University Press.

Sussman, R.S., & Sedivy, J.C. (2003). The time course of processing syntactic dependencies: Evidence from eye movements. *Language and Cognitive Processes, 18,* 143-161.

Thompson, C.K., & Shapiro, L.P. (2007). Syntactic complexity in treatment of sentence deficits. *American Journal of Speech-Language Pathology, 16,* 30-42.