The Distinction Between Unaccusative and Unergative Verbs in Turkish: An Offline and an Eye Tracking Study of Split Intransitivity

Deniz Zeyrek (dezeyrek@metu.edu.tr)   **Cengiz Acartürk (acarturk@metu.edu.tr)**

Cognitive Science, Informatics Institute
Middle East Technical University, 06800, Ankara, Turkey

Abstract

The Unaccusativity Hypothesis (UH) holds that intransitive verbs are divided into two broad classes, namely unaccusatives and unergatives. While there is evidence that the UH holds cross-linguistically, it is known that languages do not divide the intransitives into two uniform groups. We investigate the unaccusative-unergative distinction in Turkish by an offline grammaticality judgment task using a visual analog scale and by running an eye tracking experiment to tap on cognitive processing of split intransitivity. Cluster analyses indicate that the results of two experiments are broadly compatible, i.e., native speakers represent intransitive verbs in two classes, as the UH predicts. However, the offline experiment results specify uncontrolled process verbs as unaccusative, whereas the eye-gaze data characterize them as unergative. This result lends partial support for Auxiliary Selection Hierarchy. We also suggest that the uncontrolled process verb class might be where the unaccusative-unergative split occurs in Turkish.

Keywords: split intransitivity, unaccusative-unergative distinction, Unaccusativity Hypothesis, Auxiliary Selection Hierarchy, eye-tracking, reading, processing

Introduction

In modern syntax, intuitive judgments of native speakers are of great importance in empirically testing the validity of a linguistic theory. Yet, it is known that even carefully elicited (offline) grammaticality judgments give variable results because factors like the multi-dimensionality of sentence acceptability, informants' sophistication, the linguistic context, etc. interact (Gerken & Beaver, 1986). With the advent of online methods that can reveal aspects of language processing, psycholinguists and experimental linguists can have better grasp over their data, for example by comparing the results obtained from offline grammaticality judgment tests with those from online psycholinguistic tasks. The eye-tracking method gives particularly good results in this endeavor because in natural and unconstrained settings eye movements and visual attention are largely coupled (Findlay & Gilchrist, 1998). Eye movements can capture the cognitive processing of the linguistic stimuli and in turn allow researchers to compare eye-gaze data with the results from offline methods.

In this paper, we are interested in investigating a linguistic phenomenon commonly known as the unaccusative-unergative distinction (alternatively Split Intransitivity, SI). The SI is expected to exist cross-linguistically and has been shown to exist in various western languages. Typologically different languages are less investigated, though for example, works on Japanese (e.g. Kishimoto 1996, Hirakawa 1999, Sorace & Shomura 2001) and Urdu (Ahmet 2010) exist. Therefore, the current paper investigates yet another non-Indo-European language, Turkish, which is understudied with respect to the SI both linguistically and from the processing side. We first examine whether native speakers differentiate between unaccusative and unergative verbs via an offline test of grammaticality judgment. Then, to capture the real-time processing problems of native speakers with unaccusative and unergative verbs, we run an eye-tracking experiment.

Split Intransitivity and some diagnostics

Ever since the seminal work of Perlmutter (1978), the Unaccusativity Hypothesis (UH) has been widely accepted in linguistics. The UH states that, cross-linguistically, intransitive verbs fall into two broad categories with respect to their syntactic behavior, i.e. unergatives (*run, swim, talk*) and unaccusatives (*arrive, sink, bloom*). The SI rests on the idea that while the single argument of unergative verbs behaves like an agent, the single argument of unaccusative verbs behaves like an underlying theme.

Various syntactic diagnostics have been proposed to test the SI. Among the well-known ones are perfect auxiliary selection (for German, Italian, Dutch), the resultative construction (for English), impersonal passivization (for German) (Alexiadou, et al. 2004, Aranovich, 2007), and genitive negation (for Russian) (Pesetsky, 1982). In languages that have two perfect auxiliaries (equivalents of *have* and *be* in English), *have* broadly singles out unergatives, while *be* picks out unaccusatives (*German Rose hat gearbeitet; Rose ist gekommen*). In English, the resultative construction is grammatical with transitives and unaccusatives but not with unergatives (*Jane hammered the metal flat*; *The river froze solid; *Tom shouted hoarse*) (Levin & Rappaport-Hovav 1995). Yet another syntactic diagnostics is participial constructions (Kaufmann, 1995), which seems to be one of the most reliable ones across languages (Zaenen, 1993). In Turkish, a productive participial construction exists, i.e. the attributive use of the past tense suffix –*mIş* (henceforth the –*mIş* participle), e.g., *çürü-müş elma* 'the rotten apple'. This construction appears to be compatible with most unaccusatives but not with unergatives (**sıçra-mış sporcu* 'the jumped sportman'). (See Acartürk & Zeyrek, 2010 for a list of syntactic diagnostics of the SI in Turkish).

Finally, impersonal passivization (IP) is to pick out unergatives only, leading to ungrammaticality with

unaccusatives (German *Es wurde dauernd geredet; *Es wurde schnell entkommen*) (Hoekstra & Mulder 1990, Keller 2000). In Turkish, the IP is sensitive to human agency; therefore, it is grammatical with intransitives to the extent they can be construed as having a human agent. For instance, in *yan-ıl-dı* 'it was burnt' the verb is understood as having a non-specific human subject but in **yet-il-di* 'it was sufficed', since the verb cannot be construed as having a human subject, ungrammaticality results.

Research has shown that classifying intransitives as either unaccusative or unergatives is problematic because verbs belonging to one or the other category may deviate and may show variant behavior with respect to certain diagnostics across languages. Sorace (2000) argues for an Auxiliary Selection Hierarchy (ASH), arguing for the centrality of the semantic notions of agency and telicity for unaccusativity in Romance and German languages. The ASH predicts that there is (putatively universal) *gradience* among groups of intransitive verbs; i.e., while verbs belonging to certain classes (placed on the extreme positions in the hierarchy) categorically show unaccusative or unergative syntax, the verbs belonging to verb classes in between the extremes are vulnerable to deviant syntactic behavior. The extreme positions are occupied with maximally telic and dynamic verb classes (selecting *be*) and maximally agentive and nondynamic activities (selecting *have*). The verb categories of the ASH are represented in Table 1, with the categorically ('core') unaccusative verb class at the top, the categorically unergative class at the bottom.

Table 1: The verb categories of the ASH

Verb class	Example
Change of location (CoL)	arrive
Change of state (CoS)	wilt
Continuation of a pre-existing state (Sta)	survive
Existence of state (Sta)	exist
Uncontrolled process (Unc)	shine
Controlled motional process (CMP)	walk
Controlled nonmotional process (CnMP)	talk

Experimental Evidence for the Syntactic Distinction between Unaccusative and Unergative Verbs The areas of language acquisition and language attrition are usually regarded as test beds for linguistic theories. Through offline tests, a number of experimental studies in second language acquisition (L2) and language attrition have pointed out the difficulties of non-native speakers in producing unaccusatives and unergatives with appropriate syntactic constructions. For example, Sorace (1993a, b) shows that L2 learners of Italian initially exhibit consistent intuitions with respect to core unaccusative and unergative verbs, although they cannot attain native-like proficiency. Montrul (2005) discovers that bilingual heritage speakers of Spanish display an advantage over L2 speakers in syntactically distinguishing between unaccusative and unergatives verbs. While Sorace's L2 studies have generally shown that the SI is reflected in the learners' developing grammar, Montrul's

heritage language speakers provide stronger evidence for the SI, suggesting that even in language loss situations, the SI is invariant.

Studies showing the psycholinguistic reality of the unaccusative and unergative verbs also exist. In an online processing experiment, Friedmann et al. (2008) use a cross modal lexical priming technique, and find that subjects of unaccusatives reactivate after the unaccusative verb (i.e. a priming effect was found after the unaccusative verb), while subjects of unergatives do not. This result supports the idea that the single argument of unaccusatives are generated as a direct object and moved to the subject position in the syntactic tree. In a neuroimaging study, Shetreet et al. (2009) examine the cortical locations related with the comprehension of unaccusatives and unergatives and find different patterns of activations in the brain, leading them to conclude that the brain distinguishes between unaccusatives and unergatives. Their results also reveal differential cortical activities associated with syntactic and lexical operations that derive unaccusatives: the inferior frontal gyrus may be associated with the execution of the syntactic operation (where the syntactic movement of an underlying object to subject position takes place), and the middle temporal gyrus may be associated with the lexical operations (as seems to be the case in the derivation of some Hebrew unaccusatives). Lee & Thompson (2011) compare healthy subjects and agrammatic aphasia patients. They examine real-time production of unaccusatives and unergatives using an "eye-tracking while speaking" paradigm. The eye movement data reveal that for agrammatic patients, the unaccusative-unergative distinction plays a role in initial planning stage of sentences. The results lend support for the fact that human sentence production system differentially processes unaccusatives vs. unergatives, and that the distinction is preserved in agrammatic patients, though their time course of sentence planning appears to be different than healthy subjects.

Experiment 1

In Experiment 1, we elicited grammaticality judgments for the *–mIş* participle and the IP with the six verb classes in the ASH. We collapsed continuation of pre-existing state verbs and existence of state verbs under the stative class, having six verb classes to test. A set of sentences was constructed for the verbs in the *–mIş* participle and the IP to serve as the stimuli. The sentences with the CoL verbs were constructed with a goal phrase, e.g. *eve gitmiş çocuk* 'the home gone child'. A comparison of the judgments showed that when a goal phrase was used, a higher mean acceptability score was obtained than when it was not used, $t(66) = 3.35$, $p < .05$. We expected the *–mIş* participle to broadly identify unaccusative verbs (CoL, CoS, Sta); the IP to broadly identify unergative verbs (Unc, CMP, CnMP).

Participants, Materials and Design Seventy-two participants from the Middle East Technical University (METU), Turkey, participated in Experiment 1 (mean age = 28.0, SD = 6.63; 37 female). All participants were native

speakers of Turkish. For each of the six verb classes, five to nine verbs were selected. An additional set of sentences were constructed with –*mIş* for CoL verbs with a goal phrase. Impersonal passives were also formed with the same verbs. The participants were presented a total of 150 sentences (72 test sentences and 78 filler sentences). The experiment had a within-subject design with two factors: (1) syntactic diagnostic with two conditions, namely the –*mIş* participle and the IP, (2) verb class with six conditions, namely CoL, CoS, Sta, Unc, CMP, and CnMP. The dependent variable was the mean judgment scores of the participants. The judgment scores were recorded by a web-based interface designed for eliciting linguistic judgments. The participants reported their judgment scores with reference to a reference sentence by clicking on a visual analog scale, which had the end anchors "0" and "100" at the left and the right ends (a visual analog scale is a horizontal or vertical straight line used for measuring subjective judgments; see Cowart, 1997, for the use of scaling methods in eliciting linguistic judgments). The presentation order of the experimental material was randomized. The experiment was conducted in single sessions with five-minute breaks between three randomized blocks where the stimuli were presented. The duration of the experiment was approximately 35 minutes. This is a modified form of the magnitude estimation (ME) task, which we chose to use since we wanted to obtain accurate ratings (but see Sprouse 2011, who argues that the results of ME experiments are equally well informative as other types of acceptability judgment tasks).

Results Data from one participant were excluded due to a technical failure during the experiment. Data from another participant were excluded because she was a linguist (by self-assessment). Data from three participants were discarded because they did not complete the test. The remaining data from 67 participants were included in the analysis. The mean judgment scores were calculated for each test sentence, and the means for the verb classes were determined by averaging over the test sentences in each verb class, separately for the two syntactic diagnostics. An analysis of variance test was conducted to evaluate the effect of two within-subject factors on the judgment scores. We used an alpha level of .05 for all statistical tests. The results revealed a main effect for the verb class $\Lambda = .29$, $F(5, 62) = 29.6$, $p < .01$, partial $\eta^2 = .63$ and a main effect for the diagnostics, $\Lambda = .37$, $F(1, 66) = 114.41$, $p < .05$, partial $\eta^2 = .70$. Further investigations of the results for the –*mIş* participle and the IP are discussed below.

The –mIş participle The follow-up comparisons of the ANOVA test revealed significant differences in the mean judgement scores of the verbs with the –*mIş* participle. The results revealed highest judgments for the CoS verbs, lowest for the CnMP verbs. The results are shown schematically as follows. In the following notation, the greater-than (">") symbol shows a statistically significant difference between the two sides of the symbol. The "≥" symbol shows that

both statistically significant results and non-significant results are included in the two sides of the symbol. CoS received the highest mean judgment score, whereas CnMP received the lowest mean judgment score. No significant difference was observed between CoL and Sta. Finally, CoL received significantly higher scores than Unc, whereas the difference between Sta and Unc was not significant.

$$\text{CoS} > (\text{CoL} = \text{Sta}) \geq \text{Unc} > \text{CMP} > \text{CnMP}$$

For identifying the split between unaccusative verb classes and unergative verb classes in terms of the mean judgment scores, a cluster analysis was run on six variables, each corresponding to a verb class. A hierarchical cluster analysis using Ward's method produced two clusters, between which the variables were significantly different in the main. The first cluster consisted of CoL, CoS and Unc, whereas the second cluster consisted of Sta, CMP and CnMP. This result suggests that the SI in Turkish can be observed at the border between the Unc class and the Sta class, when acceptability scores for the –*mIş* participle are employed for the cluster analysis (Figure 1).
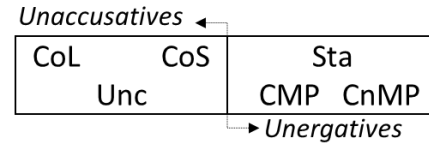

Figure 1: Split Intransitivity suggested by the results of Experiment 1.

Impersonal Passivization A comparison of the mean judgment scores among the six verbs classes revealed that the judgments for CoL verbs had highest mean scores with the IP, whereas the judgments for CoS and Sta verbs received the lowest mean scores. The results can be shown schematically below. The hierarchy is similar to the one revealed by the –*mIş* participle diagnostics, with the exception of the location of CoL.

$$(\text{CoS} = \text{Sta}) < \text{Unc} < (\text{CMP} = \text{CnMP}) < \text{CoL}$$

For the IP judgments, a cluster analysis using Ward's method revealed a slightly different picture from the –*mIş* participle: A hierarchical cluster analysis produced two clusters with significant differences between the variables in the main. The first cluster was formed by CoS and Sta, whereas the remaining four verb classes formed the second cluster. The unexpected member of the second cluster was CoL, which is assumed to be unaccusative in many languages. This behaviour of the CoL verbs may be explained by the fact that in Turkish, the IP controls for the semantic notion of agency.

To sum up, the results of Experiment 1 unambiguously showed that the –*mIş* participle places CoS in the unaccusative group and CnMP in the unergative group. A cluster analysis divides the verb classes into two parts,

placing CoS, CoL and UnC in one cluster, and Sta, CMP and CnMP in the other. On the other hand, the IP unambiguously places CoS and Sta in the unaccusative group and CoL in the unergative group. A clustering analysis for the IP produced two clusters, placing CoS and Sta in one of the clusters and the remaining four verb classes into the other cluster. The most promising finding was the results of the cluster analyses, which divided the six verb classes into two groups with respect to both diagnostics. Except for the CoL verbs which cause difficulty for the IP, the results of the cluster analyses are well-aligned, such that a split was found among the intransitives, separating CoS and Sta verbs (as unaccusatives) from the rest. This split was particularly evident with the *–mIş* diagnostics.

The results obtained in Experiment 1 are based on the mean acceptability judgments of the participants. In such experiments, participants report a judgment for the whole sentence rather than for a constituent of the sentence, but the results are usually interpreted in terms of the constituents rather than the whole sentences. This method of measuring participants' linguistic intuitions is usually called "offline" because a judgment is reported after the processing of the linguistic constituents. For this reason, the interpretation of the results obtained by offline measurement techniques is usually subject to an assumption about the cause of the difficulty during the course of processing (e.g., the resulting judgment score is the score given to the whole sentence, whereas the authors usually discuss the constituents of the sentence in isolation). This gap between the results and their interpretation can be partially closed by using online processing measures, such as the eye tracking methodology (Rayner, 1998, 2009). In this methodology, certain properties of eye movements are used as an indication of the processing difficulties a reader may experience while she is reading sentential constituents.

The frequently used eye movement parameters in online measurement of language processing are fixation duration (the duration of a single fixation on a certain location), gaze time (the sum of the fixation durations on a certain region), fixation count (the number of fixations on a certain region), and gaze regression (the number of passes on a certain region, measured in terms of fixation counts). Accordingly, longer fixation duration, longer gaze time, higher number of fixation counts, and higher number of regressions on a certain part of the text are used as measures of processing difficulty about the relevant sentential constituent. Due to its potential to reveal processing difficulties in reading, the eye tracking methodology can be used as a measure of online language comprehension processes (Rayner, 1998, 2009; Pickering, et al., 2004; see Staub & Rayner, 2007, for a review). Recently, the eye-tracking methodology has been used to identify the processing correlates of the ASH in Italian (Bard et al., 2010).

Experiment 2

In Experiment 2, we recorded the eye movements of participants while they were reading the test sentences of Experiment 1. We employed the *–mIş* participle as the diagnostic due to its higher strength as a diagnostic of the SI. We focused on the number of regressions of the participants on the critical words (i.e., the verbs with the *–mIş* participle). We expected the participants to perform more regressions toward the critical words in case of processing difficulty.

Participants, Materials and Design Twenty-nine students from METU participated in the experiment (mean age = 21.9, SD = 3.50; 19 female). All participants were native speakers of Turkish. They were presented the same set of *–mIş* participle test sentences in Experiment 1. They were asked to assess the presented stimuli according to binary judgments (i.e., grammatical vs. ungrammatical). The participants reported their judgments by pressing one of the two keys on the keyboard, one for the grammatical response and the other for the ungrammatical one. This judgment method is different than one employed in Experiment 1. The reason for using the yes/no answering paradigm was to keep the stimuli screen as simple as possible to record higher quality eye movement data for the stimuli sentences.

The eye movements of the participants were recorded by a Tobii T120, a non-intrusive, 120 Hz eye-tracker integrated into a 17" TFT monitor with 1024x768 pixels. The spatial resolution and the accuracy of the eye-tracker were about 0.30° and 0.50° degrees respectively in the product catalogue.[1] The experiment had a within-subject design. The presentation order of the experimental material was randomized. The experiment was conducted in single sessions in the Human-Computer Interaction (HCI) Lab at METU, and it lasted approximately 15 minutes.

Results Data from one participant were not included due to a total calibration failure of the eye-tracking device. The data from the remaining 28 participants were included in the analysis. Twelve of a total of 1008 eye movement protocols (for 28 participants and 36 sentences) were excluded from the analysis due to a partial calibration failure (98.8 % of all the collected data were included in the analysis). Fixations shorter than 50 ms were also excluded from the analysis. The participants' gaze regressions to the critical word (which was always the verb with the *–mIş* participle) were calculated for all the stimuli. Table 2 below shows the mean number of gaze regressions for each verb class.

An ANOVA test revealed a main effect for the differences in mean number of regressions among verb classes, $\Lambda = .32$, $F(5, 24) = 10.1$, $p < .01$, partial $\eta^2 = .68$. Further pairwise comparisons revealed the pattern of results schematically shown below, showing that CoL and Unc verbs revealed fewer number of regressions than CoS and

---

[1] In reading experiments, gaze time data are better recorded by an eye tracker of resolution 250 Hz or more (usually, 500 Hz) compared to a 120 Hz eye tracker (Holmqvist, et al., 2011). To avoid the drawbacks of the gaze time calculations with the available eye tracker, we measured gaze regressions to the critical word, which was much easier due to the relatively large saccadic amplitude in regressions.

Sta verbs, which revealed fewer number of regressions than CnMP and CMP verbs.

$$(CoL = Unc) < (CoS = Sta) \leq CnMP < CMP$$

Table 2: Mean number of regressions for the verb classes. The numbers in parentheses show standard deviation.

Verb Class	Gaze Regression (SD)
CoL	2.02 (1.64)
CoS	2.83 (1.99)
Unc	2.67 (2.12)
Sta	2.26 (1.68)
CMP	3.43 (2.12)
CnMP	2.82 (1.89)

A cluster analysis was run on six variables, each corresponding to a verb class, to identify the split between unaccusatives and unergatives with respect to eye-gaze data. A hierarchical cluster analysis using Ward's method produced two clusters, between which the variables were significantly different in the main. The first cluster consisted of CoL and CoS, whereas the second cluster consisted of the remaining verb classes, namely Unc, Sta, CMP and CnMP. The results obtained by cluster analysis did not group Unc in the CoL-CoS cluster, contra expectations revealed by the pairwise comparisons.
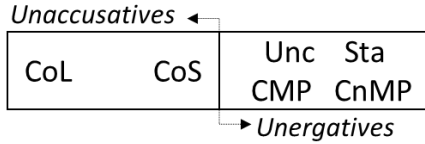

Figure 2: Split Intransitivity suggested by the results of Experiment 2.

To sum up, the results of Experiment 2 grouped CoL and CoS verbs in one cluster, the remaining verb classes in another cluster (Figure 2).

Summary, Conclusions & Future Work

In the current study we offered a description of the facts about the SI phenomenon in Turkish by an offline grammaticality judgment task and an eye-tracking experiment. The purpose of Experiment 1 was to obtain judgments from native speakers with respect to the best-known syntactic diagnostics of the SI (the *–mIş* participle and the IP) and the results informed us as to how native speakers computed contextualized intransitive verbs with these diagnostics. Experiment 2 was conducted to understand the real-time processing difficulties of Turkish native speakers while reading sentences containing *–mIş* participles with six verb classes, providing data as to how they cognitively processed the intransitive verbs.

The (offline) results of Experiment 1 (particularly those from the *–mIş* participle task) and the (online) results of Experiment 2 are broadly similar to each other. Overall, the results provide supporting evidence that there is a split among the intransitive verbs in Turkish, with the CoL and CoS verb groups as unaccusative, the CMP and CnMP groups as unergatives, in support of the UH.

The dividing line between the unaccusatives and unergatives then, might be the Unc verbs in Turkish. The Unc verbs are classified as unaccusative in the offline grammaticality judgment task and as unergatives in the eye-tracking task. The difficulty of grouping the Unc verbs with a specific intransitive class with respect to our tasks needs more explanation. As research in the last decade has shown, the variant behavior of certain verbs are a challenge for the UH. Rosen (1984) shows that many uncontrolled verbs exhibit variant behavior across languages, e.g., *sneeze* displays unergative behavior in Italian and Dutch, unaccusative behaviour in Eastern Pomo, and are compatible with both unaccusative and unergative diagnostics in Choctaw. *Bleed* fairs well with unergative diagnostics in Italian, with unaccusative ones in Eastern Pomo.

Regarding our results of Experiment 1, Sta verbs are either unaccusative or unergatives; regarding the results of Experiment 2, they are unergatives. So Sta verbs also exhibit variable behavior in Turkish, similar to many other languages. For example, in Italian *last*, in German *lay* may take either *have* or *be* (Sorace 2000).

According to Sorace (2000), variability is a notion associated with the position of a verb class on the ASH; the more a verb class is away from the extreme positions, the more likelihood it has for variant syntactic behavior. This is the sense in which she uses gradience. Verbs belonging to verb classes away from the poles of the ASH will also exhibit graded perceptions of (un)grammaticality. Both Unc and Sta are in the middle sections of the ASH, implying that they would have variable syntactic status and yield graded (un)grammaticality judgments. While we do not claim to have found gradience among the groups Turkish intransitives we looked at, we take the difficulty of categorizing the Unc and Sta verbs as unaccusative or unergative as evidence for the variability of these verb classes, in partial support of the ASH.

Some questions remain unanswered. We have not checked, for example, the collocational probabilities of the verbs with the *–mIş* participle (or the IP). Future research will show whether such statistical tendencies in the language have an effect on the acceptability judgments or eye-gaze data.

References

Acartürk, C., & Zeyrek, D. (2010). Unaccusative/unergative distinction in Turkish: A Connectionist approach. In S. Hussain, V. Sornlertlamvanich & H. Riza (Eds.), *Proceedings of the Eighth Workshop on Asian Language Resouces*. (pp. 111-119). Beijing, China: Chinese Information Processing Society of China.

Ahmet, T. (2010). The unaccusativity/unergativity distinction in Urdu. *JSAL, CSLI Publications, 3*(1), 3-21.

Alexiadou, A., Anagnostopoulou, E., & Everaert, M. (2004). *The unaccusativity puzzle: Explorations of the syntax-lexicon interface.* Oxford University Press.

Aranovich, R. (2007). *Split auxiliary systems*. Amsterdam/ Philadelphia: John Benjamins.

Bard, E. G., Frenck-Mestre, C., & Sorace, A. (2010). Processing auxiliary selection with Italian intransitive verbs. *Linguistics*, *48*(2), 325-361.

Burzio, L. (1986). *Italian syntax: A Government-Binding approach*. Dordrecht: Foris.

Cowart, W. (1997). *Experimental syntax: applying objective methods to sentence judgments*. Thousand Oaks, CA: Sage.

Gerken, L. A., & Beaver, T. G. (1986). Linguistic intuitions are the result of interactions between perceptual processes and linguistic universals. *Cognitive Science*, 10, 457-476.

Findlay, J. M., & Gilchrist, I. D. (1998). Eye guidance and visual search. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 295-312). Oxford, UK: Elsevier.

Friedmann, N., Taranto, G., Shapiro, L. & Swinney, D. 2008. The leaf fell (the leaf): the online processing of unaccusatives, *Linguistic Inquiry*, 39, 355-377.

Hirakawa, M. (1999). L2 acquisition of Japanese unaccusative verbs byspeakers of English and Chinese. In K. Kanno (Ed.), *The acquisition of Japanese as a second language* (pp. 89–113). Amsterdam: Benjamins.

Hoekstra, T., & Mulder, R. (1990). Unergatives as copular verbs: Locational and existential predication. *The Linguistic Review*, 7, 1-79.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Halszka, J., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Kaufmann, I. (1995). O-and D-Predicates: a semantic approach to the unaccusative-unergative distinction. *Journal of Semantics*, 12, 377-427.

Keller, F. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. University of Edinburgh dissertation.

Keller, F. & Sorace, A. (2003). Gradient auxiliary selection and impersonal passivization in German: An experimental investigation. *Journal of Linguistics*, 39, 57-108.

Kishimoto, H. (1996). Split intransitivity in Japanese and the Unaccusative Hypothesis. *Language*, 72, 248–286.

Lee, J. & Thompson, C. K. (2011). Real-time Production of unergative and unaccusative sentences in normal and agrammatic speakers: An eye-tracking study. *Aphasiology*, *25*(6-7), 813-825.

Levin, B., & Rappaport-Hovav, M. (1995). *Unaccusativity at the syntax-lexical semantics interface*. MIT Press.

Montrul, S. (2005). Second language acquisition and first language loss in adult early bilinguals: exploring some differences and similarities, *Second Language Research*, 21, 199-249.

Perlmutter, D. (1978). Impersonal passives and the Unaccusative Hypothesis. In *Proceedings of the Berkeley Linguistic Society*, 4, 157-189.

Pickering, M. J., Frisson, S., McElree, B. & Traxler, M. (2004). Eye movements and semantic composition. In M. Carreiras & C. Clifton (Eds.), *The online study of sentence comprehension: eyetracking, ERP, and beyond* (pp. 33-50). New York & London: Psychology Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422.

Rayner, K. (2009). Eye movements in reading: models and data. *Journal of Eye Movement Research 2*(5), 1-10.

Rosen, C. G. (1984). The interface between semantic roles and initial grammatical relations. In D. M. Perlmutter & C. Rosen (Eds.), *Studies in Relational Grammar* (pp. 38-77). University of Chicago Press.

Shetreet, E., Friedmann, N., Hadar, U. (2009). The neural correlates of linguistic distinctions: Unaccusative and Unergative verbs. *Journal of Cognitive Neuroscience*, *22*(10), 2306-2315.

Sprouse, J. (2011). A test of the cognitive assumptions of Magnitude Estimation: Commutativity does not hold for acceptability judgments. *Language*, 87(2), 274-288.

Pesetsky, D. M. (1982). Paths and categories (Doctoral dissertation, Massachusetts Institute of Technology).

Sorace, A. (1993a). Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian, *Second Language Research*, 9, 22-47.

Sorace, A. (1993b). Unaccusativity and auxiliary choice in non-native grammars of Italian and French: asymmetries and predictable indeterminancy, *Journal of French Language Studies*, 3, 71-93.

Sorace, A. (2000). Gradients in auxiliary selection with intransitive verbs. *Language*, 76, 859-890.

Sorace, A., & Shomura, Y. (2001). Lexical constraints on the acquisition of split intransitivity. Evidence from L2 Japanese. *Studies in Second Language Acquisition*, 23, 247-278.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). New York: Oxford University Press.

Zaenen, A. (1993). Unaccusativity in Dutch: Integrating syntax and lexical semantics. In J. Pustejovsky (Ed.) *Semantics and the lexicon*. Dordrecht: Kluwer Academic Publishers.