

## What indexical shift sounds like: Uyghur intonation and interpreting speech reports\*

Travis Major and Connor Mayer

University of California, Los Angeles

### 1. Introduction

Recent years have given rise to a considerable amount of research on exceptional behaviors of indexicals (e.g., I, you, here, there, etc.) in embedded contexts, a phenomenon referred to as *indexical shift* (e.g., Schlenker 2003). An example of indexical shift in Uyghur (Turkic: China, Kazakhstan, Uzbekistan) is shown in (1).

- (1) a. Tursun (men) k  t-ti-m d  -di.  
Tursun 1 SG.NOM leave-PST.DIR- 1 SG say-PST.3  
‘‘Tursun said that I<sub>Tursun</sub>/\*Speaker left.’’ / ‘‘Tursun said ‘I left.’’’
- b. Tursun \*(m  ni) k  t-ti d  -di.  
Tursun 1 SG.ACC leave-PST.DIR.3 say-PST.3  
‘‘Tursun said that I<sub>Speaker</sub>/\*Tursun left.’’

Sentences like (1a) are ambiguous as to whether they introduce direct or indirect speech (see Section 2). If (1a) is interpreted as indirect speech, the embedded 1st person, singular indexical is obligatorily interpreted as the matrix subject *Tursun*. In (1b), the accusative 1st person, singular indexical is obligatorily interpreted as the speaker of the current utterance. For present purposes, we ignore the lack of agreement on the embedded verb when the subject is accusative-marked. It should be noted, however, that regardless of agreement, both structures bear finite tense-marking and evidentiality (both cases in (1) are in the direct past).

Contrasts similar to these have been discussed in a variety of languages (e.g., Schlenker 1999, 2003, Anand and Nevins 2004, Sudo et al. 2012, Shklovsky and Sudo 2014, a.o.), but, to our knowledge, none have been studied in actual discourse contexts, or with reference to

---

\*We thank our consultants Mahire Yakup, Ziba Ablet, Mustafa Aksu, G  lnar Eziz, and Memetjan Semet for sharing their language and culture with us. Without their generosity and time, none of this would be possible. We also thank the attendees of NELS49 and UCLA Semantics Tea for feedback and conversation: in particular, Adam Royer, Beth Sturman, Hilda Koopman, Dominique Sportiche, Dylan Bumford, Ethan Poole, Byron Ahn, Harold Torrence, Gary Thoms, and Norvin Richards.

their prosodic properties. The goal of this paper is to show that considering these properties allows for a richer understanding of indexical shift. We demonstrate this by investigating indexical shift in Uyghur using a novel field methodology that involves careful control of discourse contexts and analysis of the prosodic properties of elicited utterances.

This methodology is substantially less cumbersome for consultants than the semantic diagnostics that are commonly used in the literature. It also allows for a more complete empirical picture of the phenomenon, helping to identify prosodic differences between direct quotation, indexical shift, and embedded clauses where indexical shift is blocked. This increased empirical understanding in turn informs theory: here, it suggests a necessary modification to the analysis of Uyghur indexical shift proposed by Shklovsky and Sudo (2014). Applying these methods to other languages with indexical shift will further improve both our empirical and theoretical understanding of the phenomenon.

## 2. Uyghur indexical shift

Indexical shift, first observed in Schlenker (1999), is a phenomenon where certain indexicals are interpreted relative to a context that is distinct from the current discourse context. For instance, the embedded 1st person, singular indexical in (1a) is interpreted relative to the reported utterance context where the speaker was the matrix subject *Tursun*, as opposed to the speaker of the current utterance. Although this is true if the utterance is interpreted as containing a direct quotation, such utterances may also be read as cases of indirect speech, as discussed in detail by Sudo et al. (2012) and Shklovsky and Sudo (2014).

Evidence for the existence of these indirect readings comes from Negative Polarity Items<sup>1</sup> and *wh*-questions. The *wh*-question diagnostic, which we focus on here, is based on the observation that a question embedded inside a quotation cannot be interpreted as an information-seeking, matrix scope question. For instance, the following sentence is ungrammatical in English:

- (2) \*What<sub>*t*</sub> did John say “I should eat *t*<sub>*i*</sub> tonight”?

Indexical shift, on the other hand, is a construction involves an indirect speech report that looks like quotation, but allows an embedded *wh*-expression to take matrix scope. This serves as evidence that the utterance is actually an indirect speech report where embedded indexicals are interpreted as though they were in the quotation. For instance, if we insert the *wh*-expression *qachan* ‘when’ into (1a), it can be interpreted either as a matrix scope question, or as a quoted question.

- (3) Tursun (men) qachan k et-ti-m d e-di.  
 Tursun 1SG.NOM when leave-PST-1SG say-PST.3  
 “When did Tursun say that I<sub>Tursun/\*Speaker</sub> left?” / “Tursun said ‘When did I leave?’”

<sup>1</sup>Uyghur also allows negative concord items to be licensed by negation of the matrix verb, which would not be possible in direct quotations. See Shklovsky and Sudo (2014) for more details.

### *What indexical shift sounds like*

Whether or not an utterance can be interpreted as a request for information (matrix scope wh-question) is an effective diagnostic for distinguishing indexical shift from direct quotation, because direct quotation cannot be quantified out of. The problem, however, is that quotation control, such as a matrix scope wh-question, is required for every new property that is investigated as related to indexical shift. Looking at the behavior of multiple indexicals (e.g., investigating “shift together”) in addition to the wh-question is frequently difficult for consultants.

On the other hand, indexical shift is blocked when the sentence has an accusative embedded subject and lacks agreement, as shown in (1b), repeated below as (4). This is also true of Mishar Tatar (Podobryaev 2014), Kazan Tatar (personal fieldwork), and Turkish (Deniz Özyıldız, p.c.).

- (4) Tursun \*(méni) két-ti dé-di.  
Tursun 1SG.ACC leave-PST.3 say-PST.3  
‘Tursun said that I<sub>Speaker</sub>/\*Tursun left.’

In Uyghur, there is considerable evidence that accusative subjects raise higher into the clause structure, and this is precisely how Shklovsky and Sudo (2014) account for the presence and absence of indexical shift. We revisit this account in Section 5.

### **3. Modelling of Uyghur intonation**

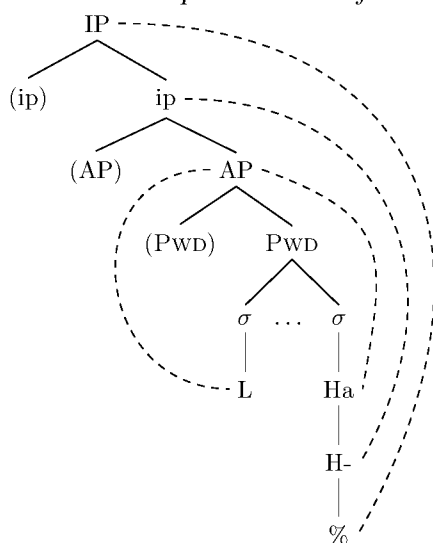
As an analytical framework in which to investigate the prosody of Uyghur indexical shift, we adopt the auto-segmental metrical (e.g., Pierrehumbert 1980, Ladd 1996/2008) model of Uyghur intonation proposed in Major and Mayer (2018). AM theory analyzes the continuous pitch contour of utterances as being realized based as a series of discrete pitch targets, consisting of high (H) tones, low (L) tones, or complex combinations of the two (e.g., HL or LH). Tonal targets associate either with prominent syllables or moras (head-marking tones), or with the left or right edges of prosodic phrases (boundary tones). The continuous phonetic pitch contour is derived by interpolation between these targets.

Uyghur exhibits only edge-marking intonation (Yakup 2013, Major and Mayer 2018). There are three levels of prosodic constituents whose edges may bear tonal targets. A schematic representation of this prosodic hierarchy is shown in (5). Solid lines indicate containment relationships<sup>2</sup>, while dotted lines reflect tonal associations. If the same edge is associated with tones from multiple prosodic levels, the highest level overrides the others.

---

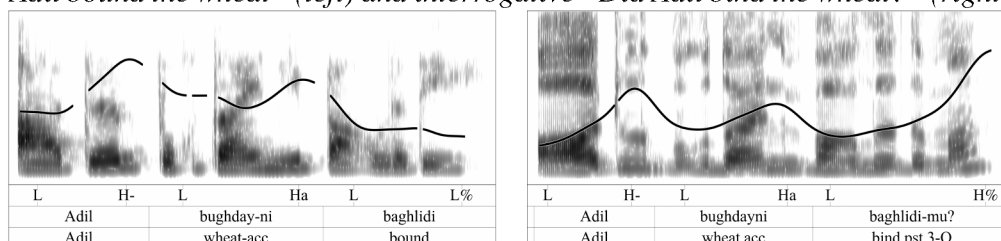
<sup>2</sup>Though we follow Ladd (1986) in allowing recursive prosodic structure as well.

- (5) A schematic representation of the Uyghur prosodic hierarchy.



The lowest prosodic level whose edges may bear tones is the *accentual phrase* (AP). An AP is marked on its left edge by a L tone, and a H target (Ha) on its right edge. APs often consist of a single prosodic word, but may also contain larger structures. The next prosodic level is the *intermediate phrase* (ip), which is marked by a high tone (H-) on its right edge. Subjects and focused constituents often form ips. Finally, the highest level is the *intonational phrase* (IP), which is marked by either a low (L%) or high (H%) tone on its right edge. These tones generally signal utterance type or continuations. H tones associated with higher prosodic levels have a higher f<sub>0</sub> than those associated with lower levels. Similarly, higher prosodic levels exhibit larger pauses at their boundaries. f<sub>0</sub> tends to trend downwards over the course of an utterance.

- (6) Spectrograms with overlaid pitch tracks and prosodic annotation for a declarative “Adil bound the wheat” (left) and interrogative “Did Adil bind the wheat?” (right).



An example of a simple declarative utterance is shown on the left in (6). The subject, *Adil*, forms an ip, while the direct object, *bughday-ni*, forms an AP. The entire utterance forms an IP with an L% final tone. A corresponding yes-no question is shown on the right in (6). The constituents are similar, but the IP now has a final H% tone.

#### 4. Stimuli and elicitation methodology

To elicit instances of direct quotation, indexical shift, and non-shifted utterances, we constructed a set of target sentences compatible with these readings. To facilitate the extraction of the pitch contour, these sentences were constructed to contain mostly voiced sounds. We then collaborated with consultants to construct short discourses containing the target sentences. For ambiguous sentences, each discourse was designed to be compatible with a reading as either a direct quotation or an embedded clause with indexical shift, but crucially not both. We recorded speakers reading through these discourses as naturally as possible.

In each case, we began by providing speakers with the precise quote that ultimately serves as the input to the reported speech. The source speech is either a declarative (7a) or a question (7b).

- (7) *Reported (a) declarative and (b) interrogative utterances compatible with indexical shift and direct quotation.*
- a. böljürgen-ni tünügün mijiwet-ti-m  
strawberry-ACC yesterday squeeze-PST-1SG  
'I squeezed the strawberries yesterday.'
  - b. böljürgen-ni qachan mijiwet-ti-m?  
strawberry-ACC when squeeze-PST-1SG  
'When did I squeeze the strawberries?'

We then provided consultants with the target sentence, such as (1a), and collaborated to build a discourse that begins with (7) and ends with the target sentence (1a). An example in English is provided below:

- (8) *A sample indexical shift discourse*

*Mahinur to me:* "I squeezed the strawberries yesterday."

*Me to you (next day):* "Mahinur said that she did a few things. She said she (lit. 'I') squeezed the strawberries."

*You to me:* "When did Mahinur say she (lit. 'I') squeezed the strawberries?"

*Me to you:* "Mahinur said she (lit. 'I') squeezed the strawberries yesterday."

This discourse contains a naturalistic production of the precise quotation, a wh-question that scopes out of the embedded clause containing shifted indexicals, and ends with the target sentence, which also exhibits indexical shift. We made sure to create multiple discourses for each construction type to avoid confounds that may arise from topic, focus, and information structure (such as the question answer pair issue above). In the case of direct quotation, we used a courtroom context where the speaker is required to provide the initial utterance verbatim. No-shift constructions were more straightforward to elicit, because there is no potential ambiguity with regard to the interpretation of indexicals.

To allow for adequate comparison between quotation and indexical shift, we used a 2x2 design that involved eliciting both declaratives and interrogatives inside both direct quotes and embedded clauses involving indexical shift. This approach was necessary so that we could directly compare cases of indexical shift where quotation control was present (matrix scope wh-question) and cases where the context enforced indexical shift, but where no diagnostic was implemented (embedded declaratives).

This elicitation strategy has a number of benefits that may not be immediately apparent. First, there is no need to elicit explicit grammaticality judgements: native speaker affirmative judgments may be assumed given that they helped to create the naturalistic context for the target sentences. In other words, the speaker is providing the context that licenses the target sentence.

Second, despite the fact that we are probing for issues related to competence, difficulty with performance can often indirectly inform us about grammaticality. For instance, if speakers continuously stumble or correct a particular piece of the stimulus sentence, it can be used as a sign that something may be problematic with the sentence and should be addressed. Finally, this method allows the researcher to identify reliable prosodic cues that discriminate between constructions. This allows for investigation of the properties of indexical shift that is more straightforward for consultants than using wh-questions (or NPI) as quotation control.

We recorded the discourses using a Zoom H4N PRO recorder and annotated them with prosodic boundaries and tonal targets using Praat (Boersma and Weenink 2019).

## 5. Results and analysis

The results of our syntactic and prosodic analysis of utterances with indexical shift, direct quotation, and no shift are schematized in (9). By tracking prosodic phrasing and tonal targets, we were able to reliably differentiate between the three constructions. We detail these properties in the following two sections.

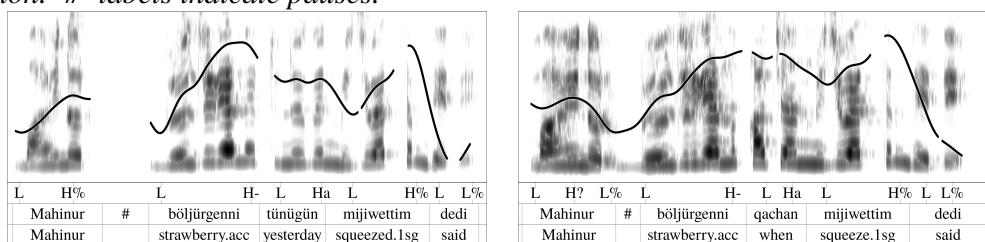
(9) *Schematic representation of the syntactic and prosodic structures of the three utterance types.*

<b>Shift</b>	[TP [Subj <sub>matrix</sub> ]	[[CP Op. [TP(Subj <sub>NOM</sub> ) V]]	V <sub>matrix</sub> ]]
	(IP (ip H-)	(ip	L/H%))
<b>Quotation</b>	[TP [Subj <sub>matrix</sub>	[[CP Op. [TP Subj <sub>NOM</sub> V	V <sub>matrix</sub> ]]
	(IP (IP L/H%)	(IP H%)	L%)
<b>No Shift</b>	[TP [Subj <sub>matrix</sub> ] [[Subj <sub>ACCi</sub>	[CP Op.[TP t <sub>i</sub> V]]	V <sub>matrix</sub> ]]
	(IP (IP L%)		L%)

### 5.1 Quotation versus indexical shift

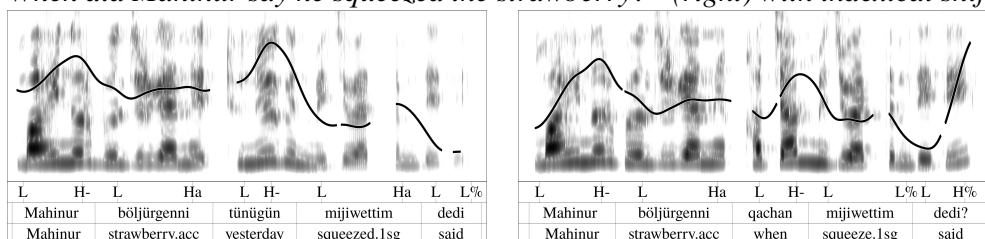
There are a number of measurable prosodic properties that effectively diagnose direct quotation. These are the consequence of the embedded clause forming its own IP (and as a result, the matrix subject also forming its own IP). First, there is consistently a large juncture at the beginning of the embedded clauses. This is true for both the declarative and interrogative embedded clauses shown in (10). Additionally, the right edge of embedded direct quotations are demarcated by a right edge H% tone, also independently of clause type. Finally, the matrix clause consistently ends with a L% tone, regardless of whether the embedded clause is declarative or interrogative. An additional, non-prosodic property that is useful as a diagnostic is that mimicry (i.e., imitation of the reported speaker) is both permitted and common in the embedded clause.

- (10) *Spectrograms with overlaid pitch tracks and prosodic annotation for a declarative “Mahinur said ‘I squeezed the strawberry yesterday’” (left) and interrogative “Mahinur said ‘When did I squeeze the strawberry?’” (right) with direct quotation. ‘#’ labels indicate pauses.*



On the other hand, utterances with indexical shift, exemplified in (11), display no evidence that the embedded clause forms an IP. This is indicated in several ways: (a) the juncture before the embedded clause is considerably smaller than in cases of quotation; (b) the matrix subject phrases with the predicate as in a simple declarative; (c) the embedded clause does not end consistently with a H% target; and (d) the clause-type tones associated with interrogatives and declaratives appear at the right edge of the matrix clause, where they do not in direct quotations. In addition, mimicry is judged as entirely unacceptable in these cases.

- (11) *Spectrograms with overlaid pitch tracks and prosodic annotation for a declarative “Mahinur said he squeezed the strawberry yesterday” (left) and interrogative “When did Mahinur say he squeezed the strawberry?” (right) with indexical shift.*

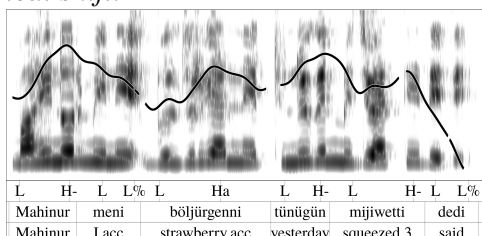


These results indicate that differentiating between quotation and indexical shift can be done effectively by simply looking at prosodic phrasing and edge tones. Following the proposal put forward in Selkirk (2011) that there is a direct mapping between syntax and prosody (‘Match Theory’), we can also conclude that the prosody of these utterances suggests that there is a closer relationship between the embedding verb and the complement clause in cases of indexical shift than in direct quotation. It appears that direct quotations, as in English, are set off almost as parentheticals, while embedded clauses involving indexical shift are prosodically treated more like regular arguments of the verb.

## 5.2 Indirect speech: No shift

Diagnosing indirect speech reports without indexical shift is fairly straightforward because it is morphologically encoded by an accusative-marked embedded subject and the absence of “subject” phi-agreement on the embedded verb (it invariably bears 3rd person). However, the prosody, exemplified in (12), does guide us toward a particular analysis of the construction. First, the subject of the embedded clause bears a L% tone on the right edge, which indicates the presence of an IP boundary. In fact, this is among the only cases we have observed in any construction where a DP does not end in a high tone, including the entire inventory of pronouns. We take this as evidence that the accusative embedded subject marks an IP boundary, associating the embedded subject with the matrix as opposed to the embedded clause. The remainder of the embedded clause is essentially indistinguishable from the embedded clause in cases of indexical shift above.

- (12) *A declarative “Adil said that I squeezed the strawberry yesterday” with no indexical shift.*



## 6. Revisiting the syntax of indexical shift in Uyghur

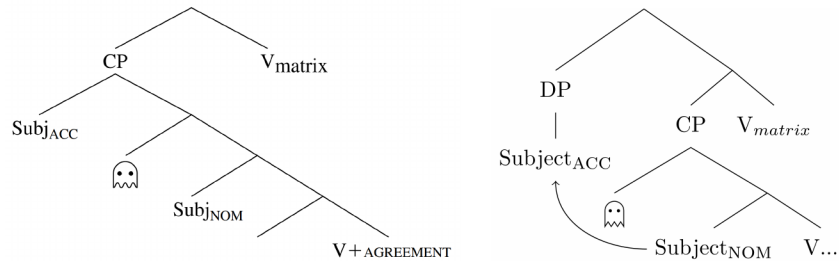
Comparing the prosody of shifted and non-shifted utterances suggests that the relationship between the verb and the complement clause it combines with is identical in both cases. This differs from the relationship between the verb and quotative clauses. The difference in prosodic phrasing of the subject of the embedded clause, however, suggests a difference in the syntax of these constructions. Shklovsky and Sudo (2014) propose the analysis shown on the left in (13). They argue that accusative subjects raise into the left periphery of the embedded clause, while nominative subjects remain inside the embedded TP. Moreover, they claim that all indexicals within the scope of the monstrous operator shift. Thus accusative subjects, which do not shift, must raise to a position outside the monster’s scope.



### What indexical shift sounds like

They provide a variety of diagnostics suggesting that the accusative subject does indeed raise, but not into the matrix clause, which we do not discuss for reasons of space.

- (13) *Analysis of Uyghur indexical shift proposed by Shklovsky and Sudo (2014) (left) and our proposed modification (right).*



The intonational data presented here suggests a necessary modification to this analysis. The accusative subject of non-shifted utterances phrases prosodically with the matrix subject, as indicated by the L% IP boundary on its right edge. The direct mapping between syntax and prosody proposed by Match Theory (Selkirk 2011) suggests that this is evidence that the accusative subject is in the matrix clause, rather than the embedded clause. We thus propose the modified analysis shown on the right in (13) (for a detailed syntactic analysis, see Major in prep.).

The difference between these two structures may seem minor, but it has important implications for the conclusions reached by Shklovsky and Sudo. There is a debate in the literature about whether the monstrous operator responsible for indexical shift is a actually a property of certain verbs themselves (e.g., Schlenker 1999, 2003) or if the operator is an independent lexical item (e.g., Anand and Nevins 2004). Shklovsky and Sudo (2014) argue that accusative subjects within the embedded clause occur in a position between the verb and the monstrous operator. In addition, they claim that because this position is higher than the position where the monster takes scope, and the semantic contribution of the verb and operator are separable, it must be the case that the context-shifting operator is an independent lexical item. Our analysis based on the prosodic properties of shifted and non-shifted utterances suggests that the accusative subject raises to the matrix clause, which no longer supports this line of reasoning. In other words, if the accusative subject actually raises into the matrix clause, there is no longer any evidence that the contribution of the matrix verb and the operator are separable – i.e., context shifting could be a verbal property.

## 7. Conclusion

This paper has shown that intonation can be a valuable field diagnostic for differentiating between quotation and indexical shift. The methodology described provides an effective way of eliciting naturalistic grammaticality judgments, while simultaneously avoiding quantificational tests involving wh-questions or NPIs, which can be extremely taxing for consultants. Using prosody as quotation control in this way can simplify tasks for speakers when investigating more complex phenomena related to indexical shift and quotation.

Furthermore, our data suggest that a modification to the analysis of Uyghur indexical shift in Shklovsky and Sudo (2014) in favor of an analysis where accusative subjects raise into the matrix clause (see Major in prep.). These prosodic results imply that the debate regarding the lexical status of the operator in Uyghur is by no means resolved. We expect that intonation will provide valuable insights into indexical shift in other languages as well.

Finally, to our knowledge this study provides the first empirical description of the the prosody of indexical shift in any language. These results provide a starting point for comparing this prosody with the prosody of related constructions that have received attention in the literature, such as parentheticals, (partial) quotation, or free indirect discourse. We expect that prosodic analysis will result in similarly valuable insights in these cases as well.

### References

- Anand, Pranav, and Andrew Nevins. 2004. Shifty operators in changing context. In *Proceedings of SALT 14*, 20–37.
- Boersma, Paul, and David Weenink. 2019. Praat: Doing phonetics by computer [Computer program]. <http://www.praat.org/>. Version 6.0.50, retrieved April 2019.
- Ladd, D. R. 1996/2008. *Intonational phonology*. Cambridge: Cambridge University Press, 1st and 2nd edition.
- Ladd, Robert D. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology* 3:311–340.
- Major, Travis. in prep. Revisiting the syntax of monsters in Uyghur .
- Major, Travis, and Connor Mayer. 2018. Towards a phonological model of Uyghur intonation. In *Proceedings of the 9th Speech Prosody International Conference*, ed. K. Klessa, J. Bachan, A. Wagner, M. Karpiński, and D. Sledziński.
- Pierrehumbert, Janet. 1980. The phonology and phonetics of English intonation. Doctoral dissertation, MIT.
- Podobryaev, Alexander. 2014. Persons, imposters, and monsters. Doctoral dissertation, Massachusetts Institute of Technology.
- Schlenker, Philippe. 1999. Propositional attitudes and indexicality. Doctoral dissertation, Massachusetts Institute of Technology.
- Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26:29–120.
- Selkirk, Elizabeth. 2011. The syntax-phonology interface. In *The Handbook of Phonological Theory*, 435–484. Oxford: Blackwell, 2nd edition.
- Shklovsky, Kirill, and Yasutada Sudo. 2014. The syntax of monsters. *Linguistic Inquiry* .
- Sudo, Yasutada, et al. 2012. On the semantics of phi features on pronouns. Doctoral dissertation, Massachusetts Institute of Technology.
- Yakup, Mahire. 2013. Acoustic correlates of lexical stress in native speakers of Uyghur and L2 learners. Doctoral dissertation, University of Kansas.