PROSODIC MARKING OF CASE AND WORD ORDER IN TURKISH SENTENCES

By

MEHMET YARKIN ERGIN

A thesis submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Dr. Karin Stromswold

And approved by

____

____

____

____

New Brunswick, New Jersey
October 2020

ABSTRACT OF THE THESIS

Prosodic marking of case and word order in Turkish sentences

By Mehmet Yarkın Ergin

Thesis Director: Dr. Karin Stromswold

A crucial step in determining the meaning of a sentence is identifying the grammatical roles of the phrases and mapping thematic roles onto them. There are (at least) five types of cues that could potentially facilitate sentence processing: propositional content, discourse context, overt case-marking, word order, and prosody. We investigated whether the way people produce spoken utterances depends on the consistency, reliability and robustness of these cues. To date, most research on spoken language production and processing has been done on languages like English that have relatively strict word order and impoverished morphology with little research on languages with flexible word order and rich morphology.

The research presented in this thesis addresses this gap by investigating the production of spoken Turkish, a language with flexible word order and rich morphology. In Turkish, sentences that have scrambled word order or lack overt object case-marking are temporarily ambiguous (i.e., garden-path sentences). We had nine Turkish speakers read aloud SOV sentences (i.e., sentences with default word order) and OVS sentences (i.e., sentences with scrambled word order) that did or did not have overt object case-marking. We found that there were prosodic differences between casemarked and non-casemarked sentences and between scrambled and non-scrambled sentences. These findings suggest that Turkish speakers prosodically mark grammatical roles when morphosyntactic cues like word order and case-marking are absent. We discuss some possible linguistic, psycholinguistic and information theoretic reasons for the observed prosodic differences, and outline future studies that could distinguish among these accounts.

ACKNOWLEDGEMENTS

I thank my advisor Karin Stromswold for her guidance and compassion through each step of this project. I also would like to thank my committee members Dave Kleinschmidt and Kristen Syrett for the valuable feedback and suggestions they provided for improving my thesis work.

Megan Kenny, Natalie Batmanian, Noyan Dokudan, Sten Knutsen, Ana Bennett, Yelda Semizer, and Elif Nur Poyraz have all helped me improve my work at various stages. I extend my gratitude to them.

I would like to express my gratitude for my family who fully supported me through this project. Without my parents and my sister, I would not be able to achieve the things that I did.

Finally, I thank my partner, Güzin Duygu Yılmaz. We supported and accompanied each other through our respective programs, in a foreign land, far from our families and friends. Without her, I would not be able to endure the emotionally challenging parts of this endeavor.

TABLE OF CONTENTS

LIST OF FIGURES

1. Introduction

1.1 Sentence processing as identifying and mapping grammatical roles to thematic roles

The cognitive tasks performed by speakers and listeners in a dialogue are remarkably different; the job of speakers is to formulate a linguistic utterance that best describes what they want to convey, and the job of listeners is to decode the utterance and understand its intended meaning. In other words, communication involves the interplay of speakers who decide how to say what they mean and listeners who interpret what the speaker meant by those utterances. A crucial step in determining the meaning of a sentence is identifying the grammatical roles of the phrases in the sentence and determining what their thematic roles are.

1.2 Grammatical & Thematic roles

Grammatical roles define formal relationships (e.g. subject, object, etc.) between a sentence's constituent phrases (e.g. noun phrases (NP), prepositional phrases (PP) etc.). Thematic roles, on the other hand, indicate how phrases function with respect to one another and to the verb they are governed by. In other words, thematic roles tell us who did what to whom in a sentence. For example, in the active sentence (1), *she* is in nominative case and, therefore, must be the subject, and *him* is in accusative case and thus must be the object. In (1), *she* has the thematic role of agent (the one doing the tipping), and *him* is the theme/patient (the one being tipped). In the passive sentence (2), the grammatical roles are flipped: *he* is the subject and *her* is a prepositional object, but the thematic roles are preserved: *he* is still the theme, and *her* is still the agent. (for more details see Chomsky, 1982)

(1) She was tipping him.

(2) He was tipped by her.

1.3 What makes sentence processing hard?

Because many sentences are temporarily ambiguous and people comprehend spoken sentences in real-time, sentence processing is notoriously difficult. Consider, for example, sentence (3) which most people need to re-read once or twice to understand[1].

(3) While the woman was dressing the baby cried. (Christianson, Hollingworth, Halliwell, & Ferreira, 2001)

The reason people need to reread this sentence is that one can interpret *dressing* as a transitive verb in which case *the baby* is an object and theme, or one can interpret *dressing* as an intransitive verb in which case *the baby* is the subject and agent of a new clause. If one assumes the former structure upon hearing *the baby*, one will be forced to reanalyze what the sentence means upon hearing *cried* because only the intransitive interpretation of *dressing* leads to a grammatically allowed structure at the end of the utterance (Frazier, 1978).

Sentences like (3), which 'lead one down the garden path,' have been traditionally called *garden-path* sentences (Bever, 1970). If people waited until the end of sentences to process them, a garden-path sentence would not pose any special difficulty. The fact that

[1] Notice how the ambiguity disappears when a comma is inserted after the *dressing*. Researchers have previously argued that the comma most likely corresponds to ortographic instantiation of a prosodic break (Staub, 2007; Breen, 2014) that makes the reader interpret *dressing* as intransitive which is the correct parsing.

people do have difficulty with garden-path sentences shows that they do not wait until the end of an utterance to parse and interpret sentences' meanings, but rather process sentences incrementally as they hear the words. (Frazier, 1978, Clifton & Frazier, 1989; Fodor, 1989; Brown, Salverda, Dilley & Tanenhaus, 2011; and for a review see Kamide, 2008 and references therein). Because sentence processing happens in real-time, people must predict what the meaningful relationships (i.e., thematic roles) are among the referents of words, based on what they have heard thus far and their prior knowledge and expections. While garden-path sentences are an extreme case of temporary ambiguity, many – if not most – sentences are temporarily ambiguous. For example, let us try to parse the example sentence (2) online. During the middle of the verb, *he was tip*--, there are (at least) two possibilities for the thematic role of *he*. If the verb continues with the passive verb participle -*ed*, *he was tipped*--, sentence (2) will be passive and *he* will have the patient (alternatively called theme) role. If the verb continues with progressive aspect -*ing*, *he was tipping*--, sentence (2) will be active and *he* will have the agent role. This means that the thematic role of *he* in sentence (2) is temporarily ambiguous.

It is impossible to maintain all parses for a sentence at any given moment simply because there are too many potential pases consistent with the sentence up to that point. Listeners can either wait until the end of the sentences to understand them, or alternatively they can process them incrementally. Given that most sentences are temporarily ambiguous, the parser must make assumptions to limit the possible search space which is infinite according to certain generative grammar theories (Chomsky, 1982). What heuristics do listeners use to allow them to process sentences in real-time while simultaneously avoiding garden-paths? What sources of information help people incrementally parse and interpret sentences correctly, and thereby avoid taking the garden path?

1.4 **Cues for incremental parsing**

1.4.1 Propositional content (what's plausible)

One possible source of information that helps us predict who is doing what to whom in a sentence is our *a priori* knowledge about the world (see, for example, Altmann & Steedman, 1988; Traxler & Tooley, 2007). Top-down information can tell us which relations are possible (or more likely) and which are not (or less likely). Correctly guessing the 'intended' meaning of the set of unordered words in (4) is easy because *dog* is the only animate entity and therefore, it must be the agent of *chew* (i.e., the one doing the chewing).

(4) {CHEW, DOG, BONE}

(5) {CHEW, BUG, DOG}

(6) {SEE, CHILD, MOTHER}

Because agents are more likely to be animate and patients are more likely to be inanimate (see the discussion on thematic hierarchy in section 1.4.4), correctly identifying the agent in (5) is slightly more difficult than (4) because *dog* and *bug* are both animate. Identifying the agent in (5) requires specific knowledge about the world, mainly that dogs sometimes can eat bugs, while generally bugs do not chew on dogs and, thus, *bug* is likely to be a patient and not an agent. Finally, accurately predicting the agent in the unordered set of words in (6) is not possible because our prior knowledge about the world by itself cannot tell us who the agent is because children can see their mothers, and mothers can see their children.

1.4.2 Discourse context as a cue for mapping

Information that a speaker and a listener share may influence the way the listener decides which interpretations of sentences are more or less likely. For example, due to the referential ambiguity of *him*, the interpretation of who is moving in (7) changes depending on the listener's prior knowledge about the actors.

(7) John and Bob are really good friends and they always help each other out. Yesterday, John called Bob about helping him move.

If the listener knows that John recently got a new apartment, the listener will assume John is calling Bob to *ask for* help. Alternatively, if the listener knows that Bob just got a new job in another city, the listener will assume that John is calling Bob to *offer* help.

1.4.3 Grammatical Case and Thematic Roles

Another source of information that may help people parse sentences correctly is overt grammatical case-marking. Case is a way of classifying nouns that reflect the nouns' dependency relationships with respect to their heads (Blake, 2001). According to some generative grammar theories (e.g., Chomsky, 1982), all NPs must receive case. However, in some languages such as English, a noun's case is not always phonologically overt. Consider, for example, sentences (8) and (9):

(8) I built the house.

(9) The house is nice.

Even though there is no way to overtly distinguish the two, *the house* is said to be in the accusative case in (8), and in the nominative case in (9).

Let us return to the garden-path sentence given in (3), which is temporarily ambiguous because *the baby* could either be the object of *dressing* or the subject of the next clause. As Christianson et al. (2001) point out, if one replaces *the baby* with an overtly case-marked pronoun, the temporary ambiguity disappears. In (10), the nominative third person singular pronoun *she* indicates that the pronoun is the subject of a new clause and not the object of *dressing*, whereas in (11), the accusative third person singular pronoun *her* indicates that the pronoun is the object of *dressing* (or a possessive pronoun of a noun phrase). In this way, overt case markers can unambiguously indicate the grammatical role played by the word that the case marker is attached to.

(10) While the woman was dressing *she* [...]

(11) While the woman was dressing *her* [...]

1.4.4 Word Order and Thematic Roles

Most languages of the world have one of the six possible word orders (SOV, SVO, OVS, OSV, VOS, VSO) as their dominant word order (Dryer, 2013). If people know what the dominant word order of their language is, they may be able to predict what the grammatical roles of words are even before they hear them. For example, given that most English sentences are SVO, English speakers might expect that the first word of a new utterance is the subject of that sentence *before* anything is actually uttered. Similarly, if OVS is the dominant word order of a language, listeners might expect the first word of an utterance to be the object of the sentence.

But how does knowing (or at least highly anticipating) words' grammatical roles (e.g., subject, object) help one understand who is doing what to whom? Starting with Fillmore (1968), many linguists have argued that there is robust mapping between thematic roles and grammatical roles. For example, Perlmutter & Postal (1984) argue that thematic roles are mapped onto grammatical roles based on a hierarchy of roles (agents before themes, themes before goals, goals before others), and that these thematic roles tend to mirror grammatical roles (agents-subjects, themes-direct objects, goals-indirect objects, etc.). In other words, phrases in the subject position tend to have the *agent* thematic role and phrases in the object position tend to have the *theme* thematic role (see Baker, 1997, for details about the 'linking problem').

In The World Atlas of Language Structures (WALS) database (Dryer, 2013), out of 1337 languages that have entries for the "Order of Subject, Object, and Verb" query, only 189 of them have no dominant word order. Furthermore, most of the world's languages have one of two dominant word orders with 565 languages having a dominant SOV word order (e.g., Turkish, West Greenlandic, Tamil, Bambara), and 488 languages having a dominant SVO word order (e.g., English, Mandarin, Indonesian, Russian). The fact that crosslinguistically, the two most common word orders are subject-initial (S-Initial) may lead people to have a subject-initial bias when they process sentences. In fact, recent research using computational models showed that communicative and cognitive constraints of speaking and listening can account for certain Greenbergian linguistic universals such as ordering of words (Greenberg, 1963; Hahn, Jurafsky & Futrell, 2020). This observation, coupled with the thematic hierarchy, may make native speakers of English *a priori* more likely to expect that the first noun they hear during an utterance will be the agent of that sentence. This expectation is born out for active sentences like (12), but not for passive sentences like (13).

(12) The child was seeing her mother.

(13) The child was seen by her mother.

Given that most sentences are in active rather than passive voice in colloquial speech, the usefulness of word order as a cue for nouns' thematic roles will depend on how strict word order is for the language in question. In strict word order languages like English where the vast majority of spoken sentences are active (Stromswold, 2006; Roland, Dick & Elman, 2007), word order is a reasonably reliable cue for predicting the grammatical and thematic roles of NPs, whereas for flexible word order languages, word order is a less reliable cue for grammatical and thematic roles.

1.4.5 Acoustic Cues

Whereas written sentences can last indefinitely and can be reread when misunderstood, spoken utterances are evanescent and cannot be "replayed" if misunderstood during the course of a conversation. On the other hand, the acoustic properties of speech (frequency, intensity, duration, prosodic breaks) can potentially convey information in a way that can not be conveyed in writing. For example, previous research shows that listeners' comprehension of sentences that are permanently ambiguous because a prepositional phrase can be attached at different possible locations in the syntactic structure is influenced by the placement of prosodic breaks in the utterance (Carlson, 2009). For example, in sentence (14), it is ambiguous whether Alex carried with one hand, or whether the robot has only one hand. In sentence (15), however, the addition of a pause between *robot* and *with* (indicated by [PAUSE]) resolves the ambiguity in favor of the former interpretation.

(14) Alex carried the robot with only one hand.

(15) Alex carried the robot [PAUSE] with one hand.

Pauses and deviations from the way an utterance is normally said may also be used to indicate what the focus of an utterance is. For example, in contrastive scenarios like (16), Birch and Clifton (2002) found that listeners judged 'the appropriateness of the intonation patterns' (Birch & Clifton, 2002, p. 574) of sentences like (16) to be better when the focus was on the novel/disambiguating parts (e.g., *bought*) rather than when the focus was on parts of the utterance that were previously established in the discourse (e.g, 'she' or 'the expensive black car').:

What did Mary do with the expensive black car?

(16) She *bought* the expensive black car.

She bought the *expensive black car*.

Many languages use prosody to indicate pragmatic focus, although the way they do so varies from language to language (Pike, 1945; Vaissière, 1995; De, Bolinger, Gibbon, Garding, T'Hart, Gronnum, Alcoba, and Murillo; 1998; Frota, 2002). Understanding this principled – but not one-to-one – relationship between the syntax of a sentence and its prosodic properties is important in order to have a full picture of spoken sentence processing.

1.4.6 Relative Strength of Cues in Different Languages

Given the differences that exist among the world's languages, it is not surprising that the robustness and reliability of cues for parsing and interpreting sentences varies among languages. For example, languages differ in word order flexibility and in how overt case-marking is. As mentioned previously, word order is a more robust cue for the grammatical and thematic roles of nominal phrases in languages like English that have a strict word order preference than in languages like Turkish that have more flexible word order. Conversely, case-marking is a more robust cue for grammatical and thematic roles in languages like Turkish that have rich overt inflectional morphology than in languages like English that have relatively impoverished overt inflectional morphology. To date, most research on spoken language production and processing has been done on languages with relatively strict word order and impoverished morphology. The research presented here seeks to address this gap by investigating the production of spoken Turkish, a language with flexible word order and rich morphology.

1.5 **Turkish**

1.5.1 Turkish Grammar

Turkish has flexible word order with SOV being the most common and pragmatically neutral way of arranging words (Erguvanli & Taylan, 1984). For example, in her analysis of child-directed speech, Batman-Ratyosyan (2003) showed that out of the 5190 adult utterances that contain a subject, object and a verb, 67% were subject-initial and 27% were object-initial, and of the utterances that start with a subject, 79% of them are either $SO_{ACC}V$ or $SO_{\emptyset}V$ sentences. [2]

Turkish has a case system with six cases: nominative, accusative, dative, locative, ablative, and genitive. In contrast to English where only personal pronouns are overtly casemarked, Turkish overtly case-marks all but subject NPs with a suffix. When a Turkish object NP is casemarked, then all of the six possible configurations are grammatical; $SO_{ACC}V$, SVO_{ACC}, $VO_{ACC}S$, VSO_{ACC}, $O_{ACC}VS$, $O_{ACC}SV$ (Kornfilt, 2013). However, when the object NP is not casemarked, only $SO_{\emptyset}V$ and $O_{\emptyset}VS$ orders are grammatical (Erguvanli & Taylan, 1984; Kural, 1992). In Turkish, casemarked objects refer to definite NPs while

[2] The folowing notational conventions are used throughout the thesis: S = Subject, V = Verb, O_{ACC} = Object with an overt accusative marker, and $O_{\emptyset}$ = Object without an overt accusative marker.

non-casemarked objects refer to indefinite NPs (Erguvanli & Zimmer, 1994) and often have an indefinite determiner (*bir*, ‘a’).

1.5.2 Turkish Garden-path Sentences

Turkish’s flexible word order means that simple declarative utterances are potentially ambiguous (i.e., garden-path sentences) with the point of morphosyntactic disambiguation differing depending on the order of the two nouns and the verb, and the presence or absence of overt accusative case-marking (Batmanian & Stromswold, 2019).

Let us first consider how overt accusative case-marking affects online intepretation of the grammatical roles of NNV sentences. Because $SO_{\emptyset}V$ sentences are grammatical in Turkish and $O_{\emptyset}SV$ are not, for NNV sentences that lack overt case-marking (i.e., $N_{\emptyset}N_{\emptyset}V$ sentences such as (a) in Table 1), the grammatical roles of the two nouns in NNV sentences become morphosyntactially disambiguated at the onset of the verb when it becomes clear that the second noun does not have an accusative casemarker. In NNV sentences that have an overt accusative casemarker (i.e., a $N_{\emptyset}N_{ACC}V$ sentence such as (b) in Table 1), the grammatical roles of the two nouns in NNV sentences becomes clear at the end of the second noun when the listener hears that the second noun has an accusative case marker. In other words, for NNV sentences, the presence of overt accusative case-marking has a rather modest impact on when the two nouns can be assigned grammatical roles, because both casemarked and non-casemarked sentences are disambiguated at the same location (Batmanian & Stromswold, 2019).

Sentence type	Example	Point of Disambiguation
a. $SO_{ø}V$ ($N_{ø}N_{ø}V$)	Kedi fare gördü. cat mouse see.PST	Onset of third word
b. $SO_{ACC}V$ ($N_{ø}N_{ACC}V$)	Kedi fareyi gördü. Cat mouse.ACC see.PST	Offset of second word
c. $O_{ø}VS$ ($N_{ø}VN_{ø}$)	Fare gördü kedi. mouse see.PST cat.	Offset of third word
d. $O_{ACC}VS$ ($N_{ACC}VN_{ø}$)	Fareyi gördü kedi. mouse.ACC see.PST cat.	Offset of first word

Table 1. Point of disambiguation for SOV and OVS sentences. Disambiguation occurs at the described constituent.

Let us now consider the online interpretation of NVN sentences. For NVN sentences that lack overt accusative case-marking (i.e., $N_{ø}VN_{ø}$ sentences such as (c) in Table 1), the grammatical role of the first noun does not become morphosyntactically disambiguated until the very end of the sentence because up until the end of the second noun there is still the possibility that the NVN sentence is an SVO sentence with overt accusative case marking (i.e., an SVO_{ACC}). This possibility cannot be ruled out until after the listener hears the entire NVN sentence and realizes that there is no accusative case marker on the second noun and, hence, the first noun must be the object of the sentence.

In striking contrast, in NVN sentences that do have overt accusative case-marking (i.e., $N_{ACC}VN_{ø}$ sentences such as (d) in Table 1), the grammatical roles of the two nouns can be determined at the end of the first noun because the first noun's accusative casemarker makes it clear that it is the object, which means that the second noun must be the subject. What this means is that, for online processing of NVN sentences, the presence of overt accusative case-marking has a large effect on when the two nouns can be assigned grammatical roles: NVN sentences with overt case-marking should pose no problems for online sentence processing, whereas NVN sentences that lack overt accusative case-marking are potentially garden-path sentences that require reanalysis (Batmanian & Stromswold, 2019).

1.5.3 Turkish Prosody

Word-level stress assignment is well-studied in Turkish (see for example, Zimmer, 1970, Sezer, 1983; Kabak & Vogel, 2001). In general, as illustrated in bold in (17), (18), and (19), when Turkish words are presented in isolation, the final syllable is perceived as stressed, regardless of the length of the word.[3]

(17)	be**re**	'hat'
(18)	be**rem**	'my hat'
(19)	bere**miz**	'our hat'

[3] An exception to this regular stress pattern are negated words, some proper names, and loanwords such as those given below.

(1)	**Git**medi.	'S/he/it did not go."
(2)	**Be**bek	'Bebek, a district in Istanbul'
(3)	Ame**ri**ka	'America'
(4)	**ga**liba	'maybe'
(5)	**pro**sedür	'procedure'
(6)	ba**da**na	'whitewash'

Less work has been done on Turkish sentential prosody than lexical prosody. In their seminal work on the intonation of pragmatically neutral Turkish sentences with overt object case marking and default word order (i.e., $SO_{ACC}V$ sentences), Ipek & Jun (2013) showed that NPs in declarative $SO_{ACC}V$ sentences follow a fairly regular low-to-high F0 increase from the first syllable to the second syllable for two syllable words and verbs in these sentences decrease in F0. If a subject NP contains more than one word, and if all the words constituting the subject NP have non-final stress when uttered in isolation (i.e. if they are polysyllabic like the words in footnote 3), then the final syllable of the final word in the NP gets a rise in F0 that is in addition to the rise in the stressed syllable that we observe when that noun is presented in isolation. If a multi-word NP-Subject has words with stress on their final syllable in isolation, then they do not show an additional rise because that word's final syllable is already expressed in its 'typical' form.

Researchers have investigated the prosody of Turkish sentences of various length that have overt morphological case-marking (Kan, 2009; Özçelik & Nagai, 2011; Özge & Bozşahin, 2010; İşsever, 2003). Özge & Bozşahin (2010) for example present a variety of 'phrasal tunes' that are associated with certain orders and with different phrase lengths. However, to date little or no work has been published that investigates the prosody of Turkish sentences that have scrambled word order or that lack overt object case-marking at the same time.

1.5.4 Sentence production as a window to sentence processing and parsing

To date, most work on sentence processing have investigated the cognitive processes that govern sentence comprehension (particularly reading comprehension), and much less attention has been paid to how sentences are produced. Much can be learned about the cognitive processes involved in sentence processing by investigating the conditions that influence how people form and produce sentences with different syntactic features (e.g., conditions that lead people to produce sentences that have particular word orders, overt case-marking, pronouns, relative clauses, etc.). Furthermore, much of the work that has been done in psycholinguistics on spoken sentence production has focused on production failures (e.g., speech errors, false starts, self-corrections and hesitations), rather than on instances where speakers successfully say what they intended to say.

One way of studying how people produce sentences is to do analyses of spontaneous speech. The problem with doing so is that, sometimes the linguistic constructions and features one wants to investigate are rarely produced (e.g., passives in English). In theory, one can get around this by studying very large corpora but, unfortunately, most of the large spoken language corpora that exist are for languages like English which have fixed word orders or are corpora that were obtained in specialized situations (e.g., the corpora found in CHILDES (MacWhinney, 2000), Pennsylvania Linguistic Consortium's (PLC) switchboard corpora (Godfrey, John & Holliman, 1993), travel agents).

One can create situations in which sentences with specific attributes are more likely to be said. For example, people are probably more likely to utter sentences with a relative clauses in situations where there are two similar entities that must be distinguished between (for example, a situation in which there are two similar-looking women, one of whom is wearing a hat). Likewise, people are probably more likely to use a pronoun to refer to an entity when the referent has already been established, and when doing so unambiguously refers to one entity (e.g., a situation where there is a man and a woman, rather than two women). Another well-known way of eliciting particular constructions is structural priming (Bock, 1986; Loebell & Bock, 2003): for example, after reading a list of sentences with a ditransitive verb such as *give* in double object format (e.g., *gave her the toy*), people are more likely to describe a similar scene with the same format compared to the prepositional format (e.g., *gave the toy to her*).

In languages with flexible word order, it may be possible to create scenarios which favor non-default word orders. For example, speakers of languages where the default word order is subject-initial might be more likely to use object-initial word orders if the object in a scenario is more important or salient than the subject. It may also be possible to create situations in which the canonical grammatical role-thematic role mapping is violated. For example, in English, passive sentences are more felicitious in sentences like *the man was killed* where the theme of an event is more salient/important than the agent, or the agent is unknown (Shintani, 1979; Nariyama 2009).

For languages in which overt case-marking is optional, it may be possible to create scenarios in which overt case-marking is more or less likely to be produced. For example, in Turkish, people might be more likely to produce sentences in which the object has overt accusative case when the scenario they are asked to describe has two NPs that are equally likely to be the subject (e.g., scenarios such as (6) which involves a man, a woman and the act of serving) than a scenario in which only one NP is plausibly the subject (e.g., scenarios such as (4) which involves a man, a meal, and the act of serving). Similarly, because as mentioned previously overtly casemarked objects in Turkish are definite and non-casemarked object are indefinite, one might be able to influence the likelihood that

Turkish-speakers provide overt case-marking by creating situations in which an entity has or has not been previously introduced in the discourse.

The problem with using techniques that favor the production of certain syntactic constructions over others is that, even the most carefully constructed scenario can only bias people to produce a particular type of sentence. The open-ended nature of the task means that, especially for uncommon constructions, participants are likely to produce sentences that are felicitious and grammatical, but not what the experimenter intended. For example, given the scenario in which there are two similar-looking women, only one of whom is wearing a hat, people could use a prepositional phrase (*the woman with the hat)* rather than a relative clause (*the woman who is wearing the hat*) to indicate which woman was being referred to.

If one is interested in studying the prosody with which people say sentences that have particular syntactic features, another experimental technique is simply to have participants read aloud sentences that have the desired feature. The reason that this is possible is that previous research has revealed that even when people read silently to themselves or when they write or type sentences, they process them in a way that suggests that they use 'implicit prosody' (Fodor, 2002). For example, consider hand-writing and typing errors such as using the word *write* instead of *right.* The error makes little sense if one only considers the orthography of the two words, but looked through the perspective of implicit speech, the error makes perfect sense. Because *right* and *write* are pronounced the same, if a person subvocalizes when they write or type sentences, this would explain why would they substitute *write* for *right*. Thus, if one is interested in studying the role of acoustic cues in language processing, another option is to have participants read aloud sentences and examine the prosodic characteristics of the sentences they produce. Doing so allows the experimenter to constrain the types of sentences people produce while eliciting sentences that are produced with relatively natural prosody.

1.5.5 Using Turkish to investigate how people assign grammatical roles to thematic roles

As discussed above, there are (at least) five types of sources of information that could potentially facilitate sentence processing: propositional content, discourse context, overt case-marking, word order, and acoustics. We hypothesize that how people produce spoken utterances will depend on the consistency, reliability and robustness of these cues. In a language like Turkish which has flexible word order and rich inflectional morphology, it would be reasonable for listeners to rely more on morphological cues than word order cues when they process sentences. We predicted therefore, that when Turkish speakers *produce* potential garden-path sentences (i.e., sentences that do not have the default word order and/or that lack overt object case-marking), they may provide prosodic cues that help their listeners avoid the garden-path effect and process the sentences correctly. In the following study, we tested this hypothesis by having Turkish speakers read aloud SOV and OVS sentences that did and did not have overt case-marking, and then analyzing the acoustic characteristics of the sentences they produced.

2. Methods

Participants. Eleven Turkish-speaking adults (six female, five male) were recruited from the student body at Rutgers University and through personal connections (family, colleagues, etc.). None of the participants had a history of speech, hearing, or language disorders. Two participants were raised in Turkish-English bilingual households, and data from these two participants were discarded from all analyses. Turkish was the first and only language learned during childhood for the remaining nine participants (5 female, 4 male). Of these nine participants, seven were between 21 and 30 years old, one was 58 years old, and one was 62 years old when they participated in the study. One was a monolingual Turkish speaker (the 58 year old woman), eight spoke English as a second language, and two spoke English and had some knowledge of a third language (German, Taiwanese) that they learned later in life.

Stimuli. As shown in Table 2, the participants read aloud four types of Turkish sentences that differed orthogonally in word order (SOV vs. OVS) and in whether or not the object received overt accusative case. Each of the 4 sentence types ($SO_{ACC}V$, $SO_{ø}V$, $O_{ACC}VS$ $O_{ø}VS$) appeared once in 36 scenarios for a total of 144 experimental sentences.

Example Scenario: “The cat saw the mouse”	Accusative Case Marking	No Case Marking
SOV	Kedi fareyi gördü. Cat mouse.ACC see.PST	Kedi fare gördü. Cat mouse see.PST
OVS	Fareyi gördü kedi. Mouse.ACC see.PST cat.	Fare gördü kedi. Mouse see.PST cat.

Table 2. Experimental stimuli.

Each of the 36 scenarios had a different transitive verb (e.g., *gör-* ‘to see’, *beğen-* ‘like’), and each verb was paired with two nouns (e.g. ‘cat- mouse’, ‘girl -boy’). Special attention was paid to ensure that the pair of nouns in each scenario were, at least *prima facie*, equally likely to be the agent or the theme of the transitive verb (e.g., cat-see-mouse, boy-like-girl). This was done to ensure the propositional content of a scenario could not be used as a cue to assign thematic roles. (For a complete list of the scenarios, see Appendix A). The first two participants received six scenarios that were subsequently judged to be somewhat unnatural by a native Turkish-speaking informant, and these six scenarios were replaced with more felicitous scenarios for subsequent participants (see Appendix B).

The 144 experimental sentences were divided into four blocks, with each block having an equal number of the four sentence types, and each of the 36 scenarios appearing only once in each block. For example, for the scenario "The cat saw the mouse," participants did not read "Cat mouse saw" and "Mouse saw cat" in the same block.

Procedure. Sentences were presented one at a time on a computer screen in random order without any discourse context. Participants were instructed to read each sentence silently until they understood its meaning, and then to read the sentence aloud in a clear and natural manner. Participants were told they could reread a sentence if they felt they made a speech error, stuttered, mispronounced a word, or were unsatisfied with how they said the sentence. If participants chose to reread a sentence, they were instructed to wait a second or two before saying the sentence again. If a participant said a sentence more than once, the version with the fewest errors and dysfluencies was analyzed. After each trial, participants hit a space bar and a central fixation cross appeared for 1225 msec followed by the next trial, with no pause between the fixation cross and the next trial. This duration was selected after piloting studies previously conducted in similar sentence processing paradigms (Knutsen, Stromswold & Kleinschmidt, 2019). After each block of 36 sentences, participants took a 5-minute break, and after they had read all four blocks of sentences, they read each block of sentences a second time. The order of the sentences within a block was different the first and the second time they read the block of sentences. Several participants said they found the sentences somewhat unnatural the first time they read them, and said that the sentences seemed more natural when they read them in the second block. For this reason, we chose to analyze the sentences said in the second half of the experiment.

The experiment was run using E-Prime 2.0 experimental software (Schneider, Eschman & Zuccolotto, 2002), and participants' utterances were recorded at 44.1 kHz using a Shure SM10 head-mounted microphone and Roland Edirol R-09 recorder. Five participants were recorded in a sound-attenuated booth, four participants were recorded in their homes in a quiet room, and one participant was recorded in a music studio. The experiment lasted approximately an hour, and participants received either course credit or monetary compensation for their participation.

Acoustic analyses. The author, who is a native Turkish speaker, hand-marked the morpheme and word boundaries in participants' audio files using the acoustic analysis software Praat (version 6.0.39, Boersma & Weenink, 2018). Once the boundaries were marked, Praat determined the duration, frequency, and intensity of each interval. As depicted in Table 3, for morpheme-level slicing, sentences without overt object case-marking were sliced into four parts (subject, object, verb stem and past tense marker), and sentences with overt accusative casemarkers were sliced into five parts (subject, object stem, accusative casemarker, verb stem and past tense marker). For word-level slicing, all utterances were sliced into three parts corresponding to the subject, the object (including the accusative marker if present), and the verb (including the past tense marker).

Sentence Type	Morpheme-Level Intervals	Word-Level Intervals
$SO_{ø}V$	Kedi/fare/gör/dü. Cat/mouse/see/PAST	Kedi/fare/gördü. Cat/mouse/see.PAST
$SO_{ACC}V$	Kedi/fare/yi/gör/dü. Cat/mouse/ACC/see/PAST	Kedi/fareyi/gördü. Cat/mouse.ACC/see.PAST
$O_{ø}VS$	Fare/gör/dü/kedi. Mouse/see/PAST/cat	Fare/gördü/kedi. Mouse/see.PAST/cat
$O_{ACC}VS$	Fare/yi/gör/dü/kedi. Mouse/ACC/see/PAST	Fareyi/gördü/kedi. Mouse.ACC/see.PAST/cat

Table 3. Morpheme and word intervals, analyzed for each sentence type.

3. Results

We hypothesized that if participants said the four sentence types in acoustically different ways, this would be evident in mean fundamental frequency (F0) and/or mean intensity of the sentences.[4] All analyses are done on normalized values of these acoustic variables. Normalization was done as follows: for each participant, the participant's overall mean F0 and mean intensity were calculated using all morpheme-level and word-level intervals from all 144 sentences (i.e., all 4 sentence types), and these overall mean values were subtracted from the mean F0 and mean intensity for each morpheme and word interval. Using each participant's normalized values, rather than their absolute values in Hz or dB allows us to better compare participants with one another because it controls for factors such as males tending to have lower F0s then females, and some participants speaking more loudly than others.

3.1. **Fundamental frequency**

3.1.1. Overall Pitch Contour of Sentences

In order to get a general sense of how word order and case-marking affects the pitch contours of Turkish sentences, we created overall pitch contours for SOV and OVS sentences (Figure 1), and pitch contours for casemarked and non-casemarked versions of OVS sentences (Figure 2a) and SOV sentences (Figure 2b), collapsing across participants and scenarios. Because Praat gives less reliable results for intervals shorter than 50 msec, we first divided each utterance into 50 msec intervals and determined the F0 for each 50 msec interval for each sentence. Normalization for these 50 msec intervals was done as

[4] We could not analyze the duration of casemarked and non-casemarked intervals because the accusative suffix means that casemarked objects are inherently longer than non-casemarked objects.

follows: for each participant, the participant's overall mean F0 and mean intensity were calculated using all 50 msec intervals, and these overall mean values were subtracted from the mean F0 and mean intensity for each 50 msec interval for that participant. Then, for each word order (Figure 1) and each sentence type (Figure 2), we calculated the mean normalized F0 for each interval for all utterances produced by all participants. All utterances were aligned at the start of the utterance, however, because participants speak at different rates, the endings were not aligned. We truncated the data after 1750 msec because less than 0.6% of trials had data beyond this point.[5]

We first explored the effect of word order on pitch contour by combining together the casemarked and non-casemarked versions of the SOV and OVS sentences (see Figure 1). Visual inspection of Figure 1 reveals that the F0s of both SOV and OVS sentences rose and then declined, with both sentences types ending at a lower F0 than they began. There were, however, differences in the pitch contours of sentences with scrambled word order (OVS sentences) and sentences with default word order (SOV sentences). Specifically, the sentence-initial F0 peak for OVS sentences (the red line in Figure 1), was higher and rose and fell more abruptly than the F0 peak for SOV sentences (the blue line in Figure 1), suggestive of the object in OVS sentences having a steeper pitch accent commonly described as H*L in the literature (Özge & Bozşahin, 2010).

[5] Because there are fewer datapoints towards the end of the utterances, the pitch contours were noisier.

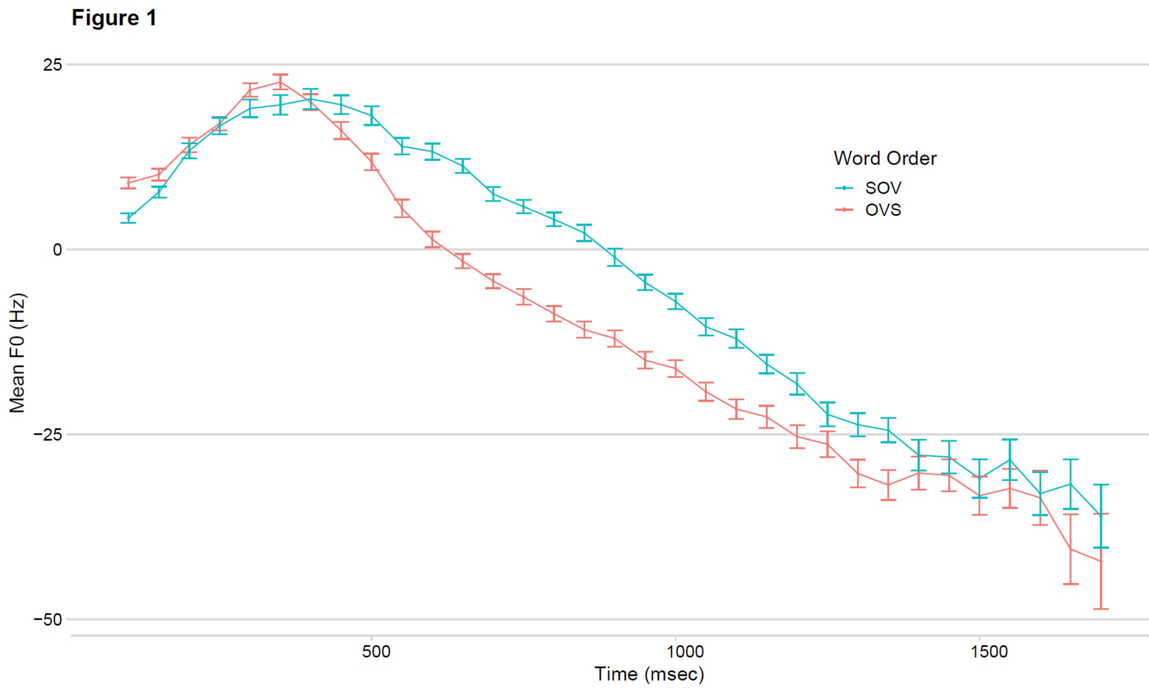


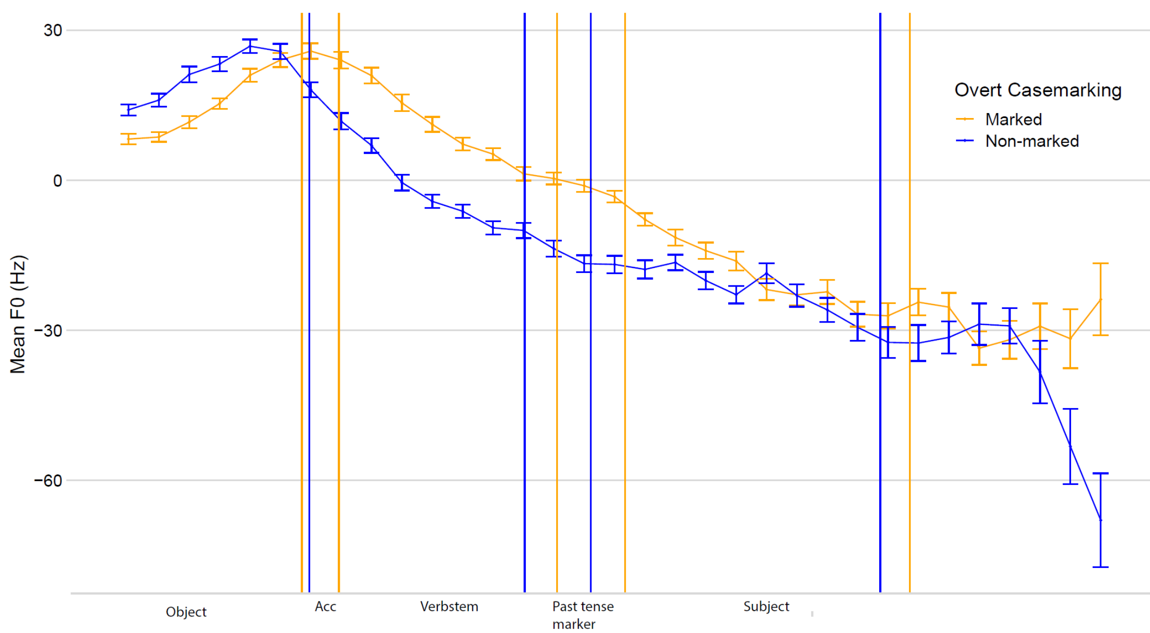
Figure 1: The effect of word order on pitch contours, collapsed across casemarked and non-casemarked sentences, participants, and scenarios. (Error bars represent standard errors)

We next explored how overt morphological case affected the pitch contour of sentences. Inspection of Figures 2a and 2b suggests that, for both OVS and SOV sentences, the pitch contours of sentences with overt case-marking (the orange colored lines in Figures 2a and 2b) differ from the pitch contours of sentences without overt case-marking (the blue colored lines in Figures 2a and 2b), with non-casemarked sentences rising more at the object stem and declining more at the verb stem than casemarked utterances did.

For OVS sentences, compared to casemarked sentences (the orange line in Figure 2a), the non-casemarked sentences (the blue line in Figure 2a) started at a higher F0, rose faster, peaked about 2/3s of the way through the object stem, and then descended more rapidly. The F0s of non-casemarked OVS sentences were lower than the F0s of casemarked OVS sentences throughout the verb stem, with the pitch contour of the two types of OVS sentence converging at the subject. Both non-casemarked SOV sentences (the blue line in Figure 2b), and casemarked SOV sentences (the orange line in Figure 2b) rose similarly during the course of the subject. At the object stem, similar to the OVS sentences, there was a crossing pattern for SOV sentences: non-casemarked and casemarked sentences began at similar F0s, but the F0 of non-casemarked SOV sentences descended more rapidly and ended at a lower F0 than the casemarked SOV sentences. At the verb stem, the separation between the two types of SOV sentences became more pronounced because the F0s of non-casemarked SOV sentences declined more rapidly than the F0s of casemarked SOV sentences.

Figure 2a

Vertical lines show mean offset.



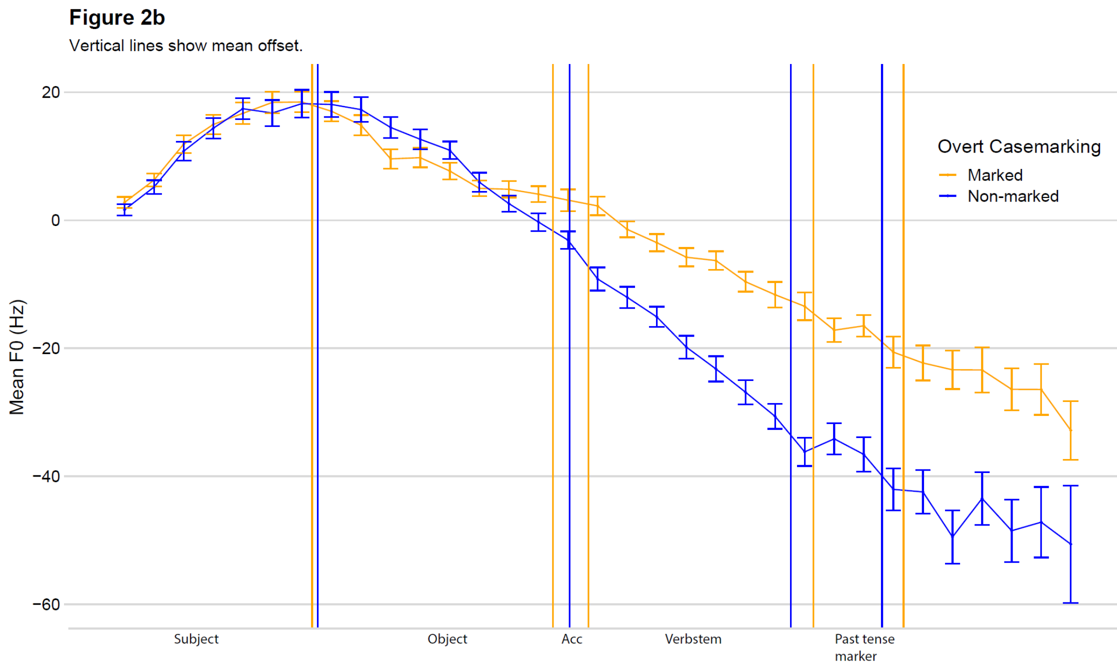


Figure 2. The effect of case-marking on pitch contours of OVS sentences (Figure 2a) and SOV sentences (Figure 2b). Error bars represent standard errors.

3.1.2. <u>Effect of Case-marking on F0 of Objects and Verbs</u>

We aligned the word and morpheme interval boundaries of the sentences and performed statistical analyses on the acoustic characterics of these intervals. All statistical analyses were performed in R (R Development Core Team, 2013) using the lme4 mixed effects linear models package (Bates, Mächler, Bolker & Walker, 2015). Sattherthwaite's method was used for approximation of degrees of freedom and *t*-tests using the same lme4 package.

For the purposes of statistical analyses, we divided each word and morpheme interval into three equal-sized parts, for a total of 9 intervals (3 each for subject, object and verb). For word-level analyses, we included the accusative markers (if present) with the objects and the past tense markers with the verbs. For the morpheme-level analyses, we divided the noun stems and verb stems into thirds, excluding the accusative markers (if present) and past tense markers. (See Table 3.)

In the first set of analyses, we analyzed the mean F0 (in Hz) of the six object and verb word-level intervals of all four sentence types.[6] We omitted the subject intervals from the analyses because their position in sentence is different for SOV and OVS (the subject is the first word for SOV and the last word for OVS), and the position of subject relative to the object is different in the two word orders (the subject comes after the object in OVS sentence and before the object in SOV). The first factor is particularly crucial because the overall decline in pitch of declarative sentences means that the subjects in SOV and OVS sentences have very different F0 profiles.

In our models, case-marking (casemarked vs non-casemarked), word order (SOV vs OVS), grammatical role (Object vs Verb), and third intervals (modeled as linear and quadratic) of the constituent were predictors along with interactions up to and including four-way terms (for details on the specification of the model's fixed effects, interactions, and error structure please see Appendix E). When a model did not converge, we removed

---

[6] Although using word-level slices causes the phonemes in the intervals of casemarked and non-casemarked objects to be somewhat misaligned, we have chosen to present the results of word-level analyses because word-level analyses work against our main hypothesis, and because noun stems in casemarked and non-casemarked object are phonetically different and have markedly different stress patterns. For example, because Turkish has word-final stress (Sezer, 1983), the final syllable of the noun stem is stressed in non-casemarked objects, whereas the (syllabic) accusative suffix is stressed in casemarked objects. This results in the final vowel of non-casemarked noun stems being longer than the final vowel in casemarked noun stems (Kabak, Maniwa & Kazanina, 2010). Even though the noun stems of casemarked and non-casemarked objects are identical in written Turkish, they differ phonetically. For example, because the vowel in the accusative suffix is always a high vowel, coarticulation causes the final vowel of casemarked noun stems to be higher than the final vowel of non-casemarked stems. The phonological differences between morpheme-level and word-level intervals not withstanding, the results of morpheme-level analyses were virtually identical to those of word-level analyses.

in step-wise fashion the random slopes accounting for the least variance until the model converged.

Overall, throughout the course of the word, the normalized mean F0s of intervals declined approximately by 9 Hz ($\hat{\beta} = -6.47$, $t(9.00) = -3.85$, $p < .005$), and most of this decline occured from the first interval to the second interval, rather than the second to the third ($\hat{\beta} = 4.25$, $t(9.09) = 4.17$, $p < .005$).

We found that, accounting for other factors, the intervals in SOV utterances were roughly 11 Hz lower than the intervals in OVS utterances ($\hat{\beta} = 5.43$, $t(160.40) = 15.65$, $p < .001$) and the normalized mean F0 of objects were higher than verbs by about 20 Hz ($\hat{\beta} = 10.22$, $t(9.34) = 7.66$, $p < .001$). These two findings probably reflect the fact that the pitch of declarative utterances tends to decline as they unfold, and the object and the verb were the first and second constituents of OVS sentences and the second and third constituents of SOV sentences. We also found that, on average, the intervals of casemarked sentences were approximately 4.5 Hz higher than the intervals of non-casemarked sentences ($\hat{\beta} = 2.21$, $t(160.31) = 6.38$, $p < .001$).

There was a two-way interaction between grammatical role and intervals. The F0s of objects increased by about 3 Hz throughout the word, while verbs decreased by 21 Hz ($\hat{\beta} = 8.97$, $t(6847.61) = 20.54$, $p < .001$). Furthermore, the decline of the verbs was slightly faster than the rise of the objects ($\hat{\beta} = -1.35$, $t(6944.99) = -3.09$, $p < .005$).

We also found a statistically significant two-way interaction between case-marking and grammatical role. Objects were about 4 Hz higher than verbs in casemarked sentences, while in non-casemarked sentences, this difference was close to 13 Hz ($\hat{\beta} = -4.17$, $t(148.12) = -13.61$, $p < .001$).

Finally, there was a significant three-way interaction among intervals, grammatical role, and case-marking. While the intervals of both casemarked and non-casemarked objects had similar F0 range (Casemarked objects: 1st interval = 0.72 Hz, 2nd interval = -3.11 Hz, 3rd interval = 2.39 Hz; Non-casemarked objects: 1st interval = -1.89 Hz, 2nd interval = -1.61 Hz, 3rd interval = 3.50 Hz), the verbs of non-casemarked sentences declined by roughly 4 Hz more than the verbs of casemarked sentences (Casemarked verbs: 1st interval = 11.89 Hz, 2nd interval = -3.82 Hz, 3rd interval = -8.07 Hz; Non-casemarked verbs: 1st interval = 14.51 Hz, 2nd interval = -5.32 Hz, 3rd interval = -9.18 Hz; $\hat{\beta} = -1.32$, $t(6944.86) = -3.03$, $p < .005$). We also observed that the difference between the rate of rise of object and decline of verb intervals was slightly larger in non-casemarked sentences than in casemarked sentences ($\hat{\beta} = 0.92$, $t(6943.80) = 2.12$, $p < .05$).

We compared the full model (i.e., four-way interaction) with a model that lacked four-way interaction terms (i.e., lacking estimates for word order interactions). Comparisons of model likelihoods showed that the model with the four-way interaction was a better fit to the data than the model without the four-way interaction.

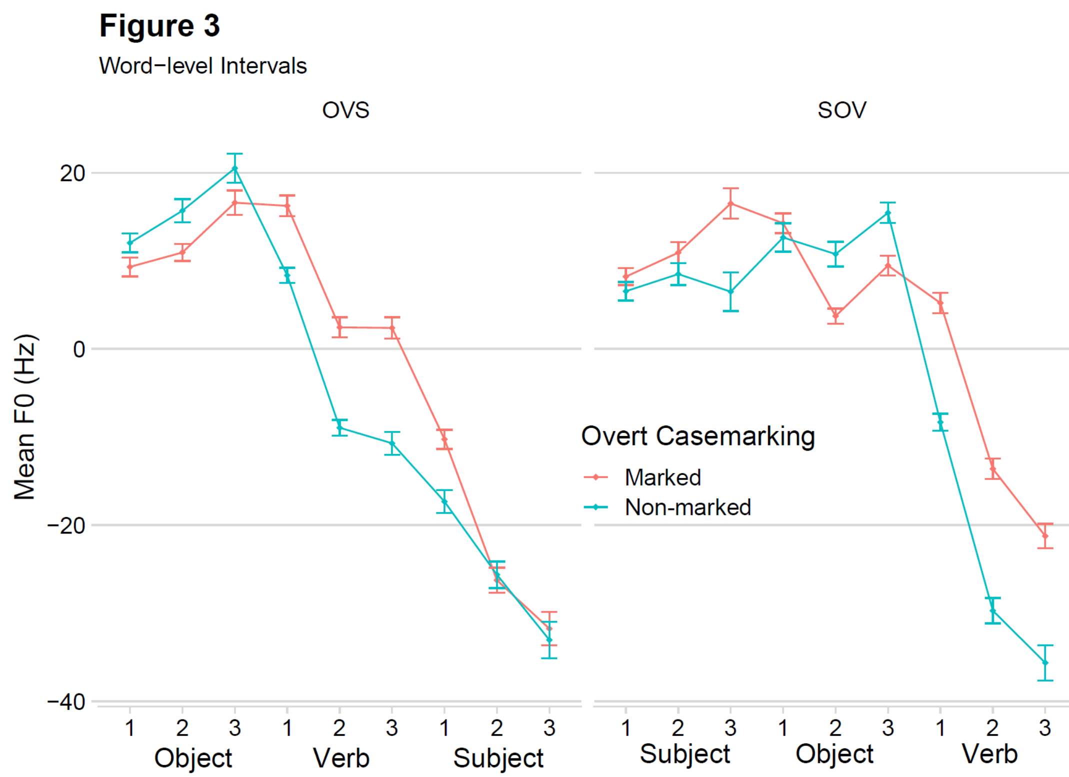


Figure 3. The effect of case-marking on F0 values for OVS sentences and SOV sentences. Bars represent standard errors.

The three-way interaction between intervals, grammatical role and case-marking can be appreciated by comparing the normalized F0 values for the object and verb intervals in OVS sentences (the left panel of Figure 3) with those in SOV sentences (the right panel in Figure 3). For OVS sentences, the F0s of all three object intervals were higher for non-casemarked sentences than casemarked sentences, and the F0s of all three verb intervals were higher for casemarked sentences than non-casemarked sentences. For SOV sentences, only the F0s of the last two object intervals were higher in non-casemarked sentences than casemarked sentences, and the F0s of all three verb intervals were higher in casemarked sentences than non-casemarked sentences. In addition, for both OVS and SOV sentences, the drop in F0 between the object and the verb was larger for non-casemarked sentences than casemarked sentences, with the difference between the F0 slopes for casemarked and noncasemarked sentences being greatest from the end of the object to the beginning of the verb.

To ensure these results were not an artifact of aggregating the data from multiple participants, we analyzed the F0 data for each participant separately. (See Appendix C for individual participants' graphs.) As shown in Table 4, for the objects in OVS sentences, 6 of the 9 participants had at least one object interval that was higher for non-casemarked than casemarked sentences, one participant had one object interval that was higher for casemarked than noncasemarked sentences, and two participants had no F0 differences for any object intervals. For the verbs in OVS sentences, 8 of the 9 participants had at least one verb interval that was lower for non-casemarked than casemarked sentences, and two participants had one verb interval that was higher for casemarked than non-casemarked sentences.

For the objects in SOV sentences, 6 of the 9 participants had at least one object interval that was higher for non-casemarked than casemarked sentences, one participant had one object interval that was higher for casemarked than non-casemarked sentences, and three participants had no F0 differences for any object intervals. For the verbs in SOV sentences, 8 of the 9 participants had at least one verb interval that was lower for non-casemarked than casemarked sentences, one participants had no F0 differences for any verb intervals, and no participants had any verb intervals that were higher for non-casemarked than noncasemarked sentences.

	OVS Sentences		SOV Sentences	
Subject	Object Intervals	Verb Intervals	Object Interval	Verb Interval

	+	-	+	-	+	-	+	-
1	3	0	0	3	2	0	0	3
2	3	0	0	2	2	1	0	1
3	1	0	0	2	2	0	0	3
4	2	0	0	2	1	0	0	3
5	0	1	0	3	1	0	0	3
6	1	0	0	3	0	0	0	3
7	3	0	0	2	0	0	0	1
8	0	0	1	0	0	0	0	2
9	0	0	1	1	2	0	0	0
Total	13	1	2	18	10	1	0	19

Table 4. Individual participants' F0 for object and verb intervals.

Key

'0' indicates that the standard error bars for casemarked and non-casemarked sentences' fundamental frequencies overlapped in all of the three intervals.

'+' indicates the number of intervals where the non-casemarked sentences had non-overlappingly higher fundamental frequences than the casemarked sentences.

'-' indicates the number of intervals where the non-casemarked sentences has non-overlappingly lower fundamental frequencies than the casemarked sentences.

3.1.3. F0 and Early Disambiguation of Word Orders

We next narrowed our focus on the F0s of the first word of the sentences in order to investigate the early effects of scrambling the word order and case-marking, keeping in mind the points of morphological disambiguation given in Table 1 and discussed in section 1.6.2 of the Introduction. We analyzed the F0 data for the first word with a mixed effects linear model that included word intervals in thirds, word order, and case-marking as predictors and all possible interactions among these three predictors. (see Appendix E for details on the model).

We found that, accounting for other factors, the intervals in SOV utterances were roughly 4 Hz lower than the intervals in OVS utterances ($\hat{\beta}$ = 2.24, $t(11.11) = 2.46$, $p < .05$) We also found a significant interaction between the third intervals and word order: on average, the intervals of OVS sentences rose more than the intervals of SOV sentences (OVS: 1st interval = -3.63 Hz, 2nd interval = -0.87 Hz, 3rd = 4.50 Hz; SOV: 1st interval = -2.24 Hz, 2nd interval = 0.13 Hz, 3rd = 2.10 Hz; $\hat{\beta}$ = 1.34, $t(3516.67) = 2.12$, $p < .05$). Although there were no meaningful difference between casemarked and non-casemarked sentences collapsed across intervals, we observed a significant two-way interaction between casemarking and intervals which showed that casemarked intervals rose more than the non-casemarked intervals (Casemarked: 1st interval = -3.46 Hz, 2nd interval = -1.14 Hz, 3rd = 4.60 Hz; Non-casemarked: 1st interval = -2.41 Hz, 2nd interval = 0.40 Hz, 3rd = 2.00 Hz; $\hat{\beta}$ = 1.29, $t(3515.99) = 2.04$, $p < .05$).

There was a two-way interaction between casemarking and word order. For OVS sentences, the F0s of non-casemarked intervals were approximately 4 Hz higher than casemarked intervals. For SOV sentences the directions was reversed: the F0 of non-casemarked intervals were approximately 4.5 Hz lower than the casemarked intervals. ($\hat{\beta}$ = -2.14, $t(148.59) = 20.18$, $p < .001$).

Finally, there was a significant three-way interaction among intervals, word order, and case-marking ($\hat{\beta}$ = -0.43, $t(3512.92) = -2.72.18$, $p < .01$) which is best appreciated through visual inspection of Figure 4 which reveals the following interplay among the three predictors. For sentences that have overt object case-marking, the pitch contours of the first words (i.e., the objects in O_iVS sentences, and the subjects in SO_iV sentences) were very similar, increasing over the course of the first word. In contrast, for non-casemarked sentences, the F0 of non-casemarked objects in O$_{ø}$VS sentences began about 5 Hz higher than the F0 of the subjects in SO$_{ø}$V sentences. In addition, the F0s rose linearly over the course of the non-casemarked objects in the O$_{ø}$VS sentences, whereas the F0s remained constant over the course of the subjects in the SO$_{ø}$V sentences.

Figure 4

Word-level intervals from only the first word.

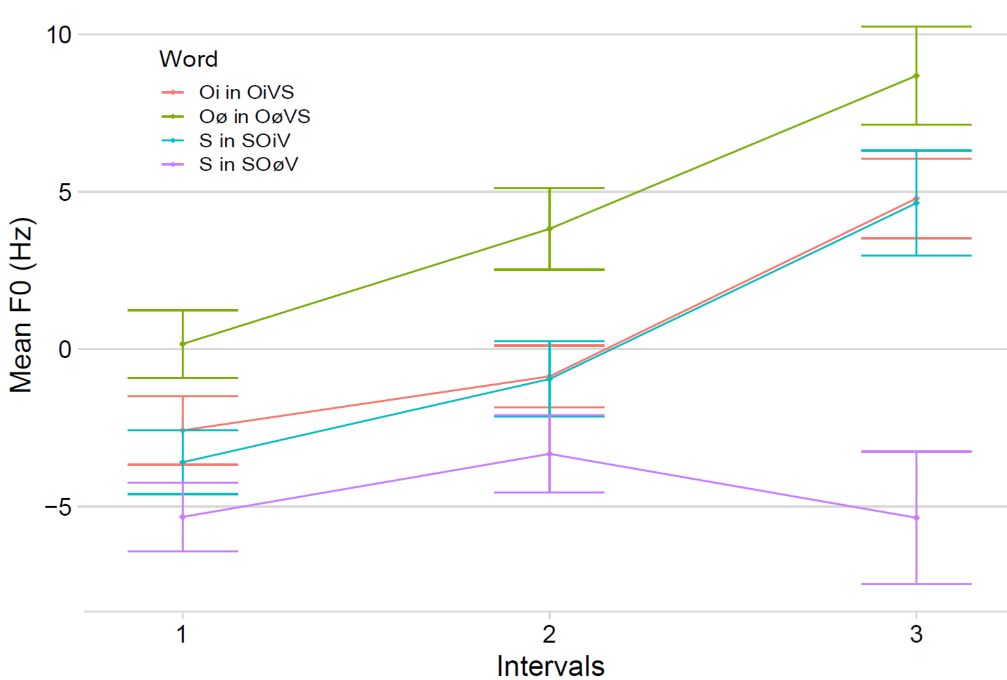


Figure 4. F0 of the first constituent in the 4 sentence types. (Error bars are standard errors.).

3.2. **Intensity**

3.2.1. Overall Intensity Contour of Sentences

We next analyzed how word order and case-marking affects the intensity (in dB) of Turkish sentences using the same techniques we used to analyze F0. We explored the effect of word order on intensity contour by combining together the casemarked and non-casemarked versions of the SOV and OVS sentences (Figure 5). Visual inspection of Figure 5 reveals that the dBs of both SOV and OVS sentences declined, with both word orders ending at a lower dB than they began. There were, however, slight differences between the intensity contours of sentences with scrambled word order (OVS sentences) and sentences with default word order (SOV sentences). Specifically, the OVS sentences declined at a fairly steady rate across all constituents, whereas the SOV sentences intensity declined early and sharply, then plateaued in the middle then declined at the end.

Figure 5

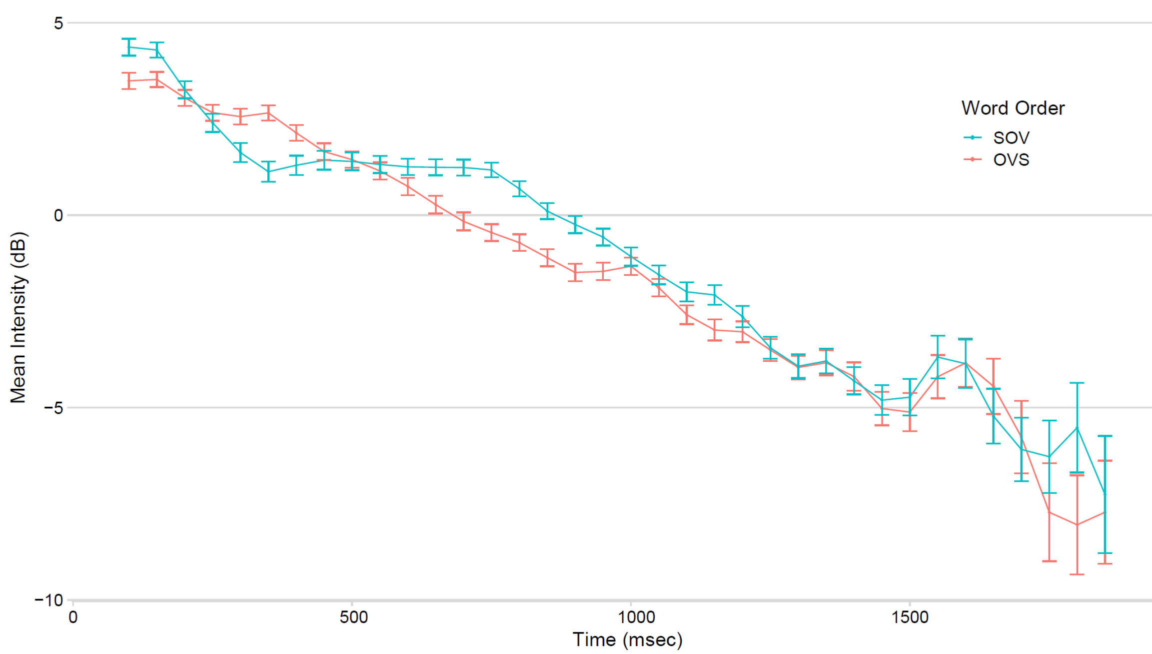
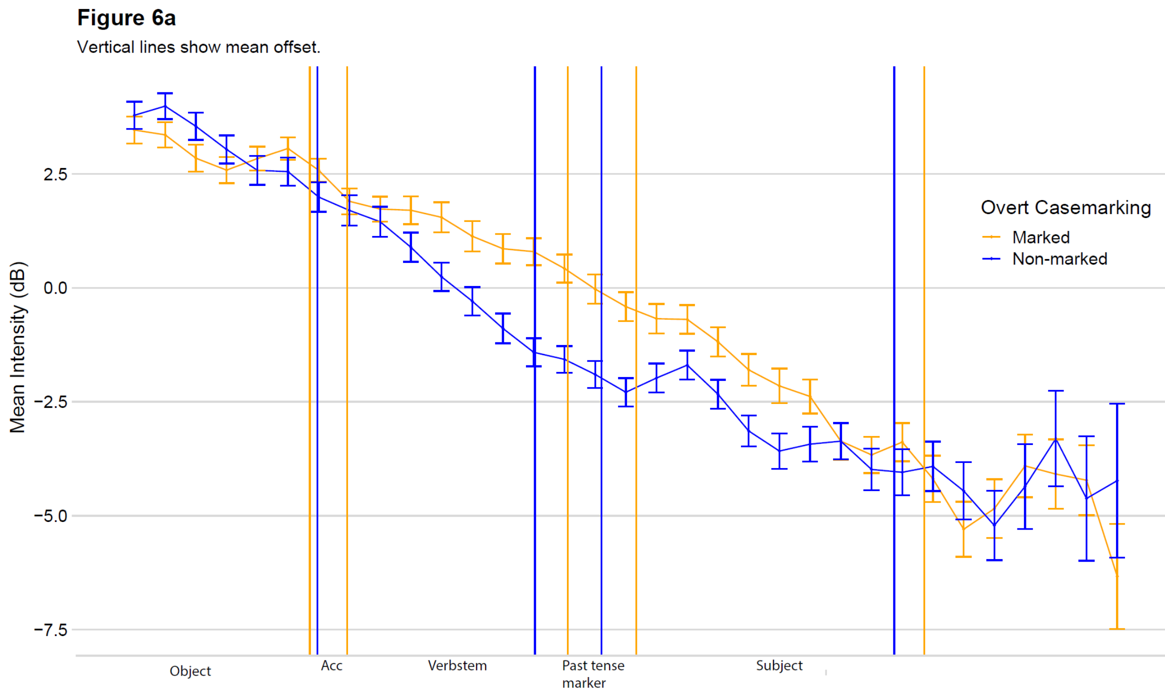


Figure 5. The effect of word order on intensity contours, collapsing across casemarked and non-casemarked sentences, participants and scenarios. (Error bars represent standard errors).

As can be seen in Figures 6a and 6b, the intensities of all four types of sentence declined monotonically, but there were differences among them. Inspection of Figures 6a and 6b suggests that, for both OVS and SOV sentences, the intensity contours of sentences with overt case-marking (the orange colored lines in Figures 6a and 6b) differ from the intensity contours of sentences without overt case-marking (the blue colored lines in Figures 6a and 6b). For non-casemarked OVS sentences, objects rose a little before declining continuously, while for casemarked OVS sentences the object stems fell before rising a little at the accusative marker. At the verb stem, the non-casemarked and casemarked OVS sentences clearly diverge, with verb stems from non-casemarked OVS sentences declining more and more rapidly than verb stems from casemarked OVS sentences. At the subject, the non-casemarked OVS sentences rose a bit before declining, whereas the casemarked OVS sentences continuously declined. The intensity curves of casemarked and non-casemarked OVS sentences roughly converged towards the end of the subject.

Figure 6b shows that the intensity of both casemarked and non-casemarked SOV sentences started high and dropped steadily throughout the subject, with the rate of decline being slightly greater for non-casemarked SOV sentences. At the object stem both casemarked and non-casemarked SOV sentences plateaued and the intensity of the two conditions were similar to each other. However at the end of the object and at the beginning of the verb stem the two types of SOV sentences began to diverge. Similar to the OVS sentences, the non-casemarked SOV sentences declined more and faster than the casemarked SOV sentences, and the intensity curves of the sentences never converged.



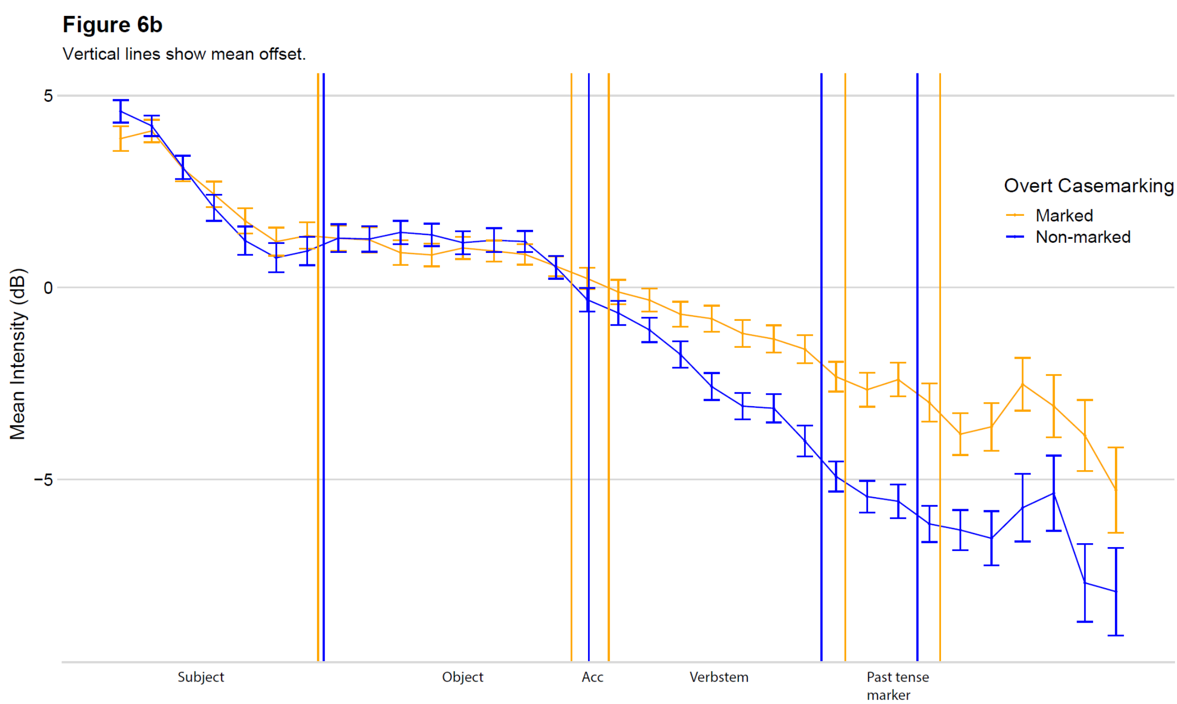


Figure 6. The effect of case-marking on intensity contours of OVS sentences (Figure 6a) and SOV sentences (Figure 6b). Error bars represent standard errors.

3.2.2. Effect of Case-marking on Intensity of Objects and Verbs

We next analyzed the normalized intensity (in dB) of object and verb intervals using the same methods that we used to analyze F0 (see section 3.1.2., and Appendix E for model tables).

Overall, throughout the course of the word, the normalized mean intensities of intervals declined approximately by 1 dB ($\hat{\beta}$ = -0.69, $t(50.27) = -3.48$, $p < .005$), and most of this decline occured from the first interval to the second interval, rather than from the second to the third ($\hat{\beta}$ = 0.46, $t(7250.99) = 6.75$, $p < .001$).

We found that, accounting for other factors, the intervals in SOV utterances were roughly 2 dB lower than the intervals in OVS utterances ($\hat{\beta}$ = 1.02, $t(174.14) = 5.50$, $p < .001$) and the normalized mean intensities of objects were higher than verbs by about 2 dB ($\hat{\beta}$ = 1.10, $t(21.35) = 4.34$, $p < .001$). These two findings probably reflect the fact that the intensity of utterances tend to decline as they unfold, and the object and the verb were the first and second constituents of OVS sentences and the second and third constituents of SOV sentences. We did not find an overall difference in intensity between the intervals of casemarked sentences and the intervals of non-casemarked sentences.

There was a two-way interaction between grammatical role and intervals. The intensities of objects remained relatively constant throughout the word, while verbs decreased by 2 dB ($\hat{\beta}$ = 0.70, $t(7251.14) = 10.32$, $p < .001$). Furthermore, the decline of the verbs was slightly faster than the slight changes between the intervals of the objects ($\hat{\beta}$ = -0.40, $t(7250.99) = -5.96$, $p < .001$).

We also found a statistically significant two-way interaction between case-marking and grammatical role. The object intervals were approximately 1 dB higher in non-casemarked sentences than in casemarked sentences, while the verb intervals were roughly

1.5 dB lower in non-casemarked than in casemarked sentences ($\hat{\beta}$ = -0.58, $t(175.19) = -3.73$, $p < .001$).

Finally, there was a significant three-way interaction among intervals, grammatical role, and case-marking. The intervals of casemarked objects decreased while the intervals of non-casemarked objects had small rise in intensity (Casemarked objects: 1st interval = 1.67 dB, 2nd interval = -0.87 dB, 3rd interval = -0.80 dB; Non-casemarked objects: 1st interval = -0.32 dB, 2nd interval = 0.12 dB, 3rd interval = 0.20 dB). Also, the verbs of non-casemarked sentences declined by roughly 1 dB more than the verbs of casemarked sentences (Casemarked verbs: 1st interval = 0.99 dB, 2nd interval = -0.53 dB, 3rd interval = -0.46 dB; Non-casemarked verbs: 1st interval = 1.67 dB, 2nd interval = -0.87 dB, 3rd interval = -0.80 dB; $\hat{\beta}$ = -0.36, $t(7251.10) = -5.33$, $p < .001$). We also observed that the difference between the rate of intensity change of object and verb intervals was slightly larger in non-casemarked sentences than in casemarked sentences ($\hat{\beta}$ = 0.21, $t(7250.98) = 3.07$, $p < .005$).

We compared the full model (i.e., four-way interaction) with a model that lacked four-way interaction terms (i.e., lacking estimates for word order interactions). Comparisons of model likelihoods showed that the model with the four-way interaction was a better fit to the data than the model without the four-way interaction.

Figure 7

Word−level Intervals

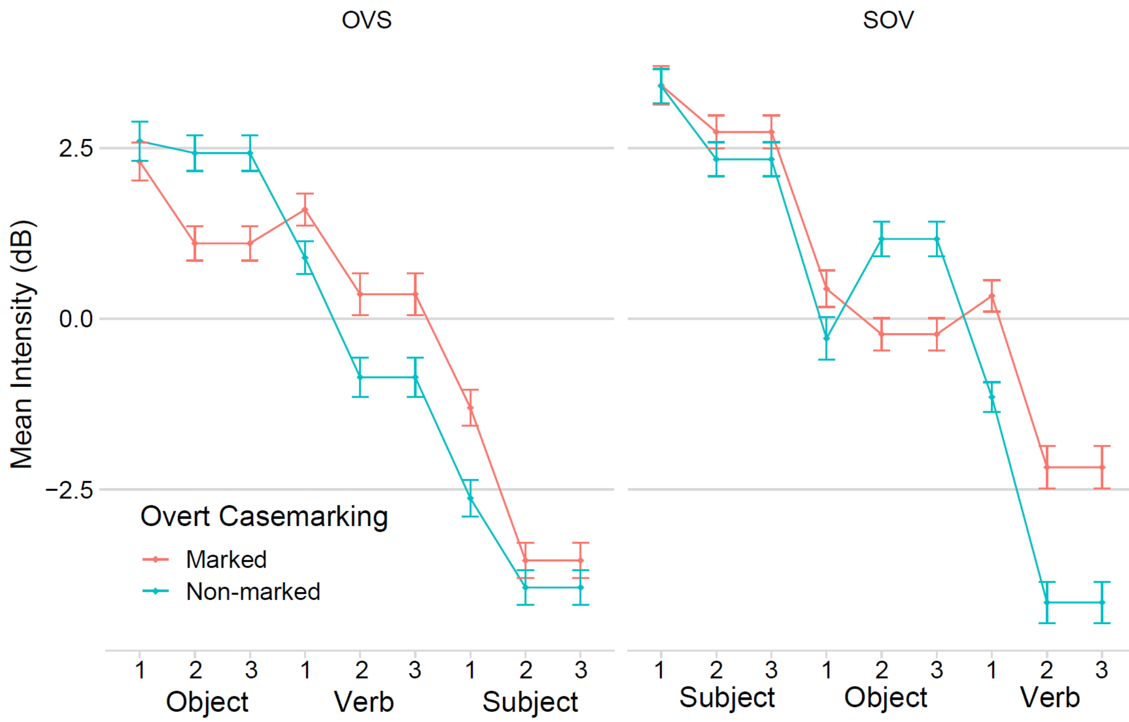


Figure 7. The effect of case-marking on dB for OVS sentences and SOV sentences. Bars represent standard errors.

The three-way interaction between intervals, grammatical role and case-marking can best be appreciated by comparing the intensities of casemarked and non-casemarked OVS sentences (Figure 7a) and SOV sentences (Figure 7b) separately. For non-casemarked OVS sentences, the intensity of the object was constant throughout all three intervals remained the same, with the last two intervals of non-casemarked objects being higher than the intensities of the last two intervals of objects in casemarked OVS sentences. The object of casemarked OVS sentence declined after the first interval, but remained steady for the last two. The intensities of all three verb intervals were higher for casemarked OVS sentences than non-casemarked OVS sentences.

For SOV sentences, the non-casemarked objects rose from the first to the second interval, and remained high at the third interval. The objects of casemarked SOV sentences declined after the first interval and remained steady, similar to what was found with OVS sentences. The intensities of all three verb intervals were higher for casemarked SOV sentences than for non-casemarked SOV sentences.

For both OVS and SOV utterances, the drop in dB between the object and the verb was larger for non-casemarked sentences than casemarked sentences, with the difference in the dB slopes for casemarked and non-casemarked sentences being greatest between the end of the object and beginning of the verb. We compared this model with the model that lacked the three-way interaction, and a χ^2 goodness-of-fit test and a comparison of AICs of the two models revealed that the model with the three-way interaction was a better fit to the data.[7]

We next analyzed the intensity of object-verb intervals for each participant separately. (See Appendix D for individual participants' graphs.) For the OVS sentences, three of the 9 participants had two object intervals that had higher intensities in non-casemarked than casemarked sentences, and six participants had no significant dB differences for any object intervals (see Table 5). For the verbs in OVS sentences, three of the 9 participants had at least two verb intervals that was lower in non-casemarked than

[7] As was the case for our F0 analyses, we compared this model with a higher order model which allows word order to interact with other predictors (i.e., 4-way interaction model). The outcome was similar to the F0 case, and due to same reasons we present the three-way model.

casemarked sentences, and six participants had no dB differences for any verb intervals. For the SOV sentences, three of the 9 participants had two object intervals that were higher in intensity in non-casemarked than casemarked sentences, and six participants had no dB differences for any object intervals. For the verbs in SOV sentences, 8 of the 9 participants had at least one verb interval that was lower in intensity in non-casemarked sentences than in casemarked sentences, and one participant had no dB differences for any verb interval. Taken as a whole, the intensity findings mirrored the F0 findings, but the pattern was less robust for dB than F0, and not all participants displayed the pattern with dB. Interestingly, participant 5, who did not show the effect for F0, and P3, who showed a smaller F0 effect than the other participants, both displayed a robust effect for dB.

Subject	OVS Sentences				SOV Sentences			
	Object Intervals		Verb Intervals		Object Interval		Verb Interval	
	+	-	+	-	+	-	+	-
1	0	0	0	2	2	0	0	3
2	0	0	0	0	0	0	0	2
3	2	0	0	0	2	0	0	2
4	0	0	0	0	0	0	0	1
5	2	0	0	3	2	0	0	3
6	0	0	0	0	0	0	0	2
7	2	0	0	2	0	0	0	1
8	0	0	0	0	0	0	0	1
9	0	0	0	0	0	0	0	0
Total	6	0	0	7	6	0	0	15

Table 5. Intensity of individual participants' object and verb intervals.

Key

'0' indicates that the standard error bars for intensity overlapped for casemarked and non-casemarked sentences in all of the three intervals.

'+' indicates the number of intervals where the non-casemarked sentences had non-overlappingly higher intensities than the casemarked sentences. '-' indicates the number of intervals where the non-casemarked sentences has non-overlappingly lower intensities than the casemarked sentences.

3.2.3. Intensity and Early Disambiguation of Word Orders

As we did for F0, we investigated the early effects of word order scrambling by comparing the intensity of the first words of non-casemarked sentences (i.e., comparing the S in $SO_øV$ sentences with the O in $O_øVS$ sentences), and the first words of casemarked sentences (i.e., comparing the S in SO_iV sentences with the O in O_iVS sentences), keeping in mind the points of disambiguation discussed in section 1.6.2. and in Table 1.

Overall, throughout the course of the word, the normalized mean intensities of intervals declined approximately by 0.75 dB ($\hat{\beta} = -0.54$, $t(9.02) = -4.45$, $p < .005$), and most of this decline occured from the first interval to the second interval, rather than the second to the third ($\hat{\beta} = 0.31$, $t(3683.59) = 3.08$, $p < .005$). None of the estimates for fixed effects or interactions had enough certainty to reach statistical significance, except for a single 3-way interaction between intervals, casemarking, and word order ($\hat{\beta} = -0.43$, $t(3512.92) = -2.72.18$, $p < .01$) which is best interpreted through visual inspection of Figure 8, which depicts the intensity of the first noun for the four sentence types. For sentences that have overt case-marking (red and blue lines in Figure 8), the intensity of the noun decreased after the first interval, and all of the intervals of the subject in SO_iV sentences were higher than the corresponding intervals in the object in O_iVS sentences. For non-casemarked sentences (green and purple lines in Figure 8), subjects in $SO_øV$ began at a higher intensity than objects in $O_øVS$ sentences, but after the first interval, the intensity of the subject in

SO$_{\o}$V sentences dropped, whereas the intensity of the object in O$_{\o}$VS sentences remained constant. The result was that only the first interval of the subject in SO$_{\o}$V utterances was higher than the corresponding interval in the object in O$_{\o}$VS utterances. Although the intensities of the four sentence types differed enough to reach statistical significance, the effect observed for dB was not the same type or direction as that observed in the F0 analysis of the first words.

Figure 8

Word−level Intervals from only the first word.

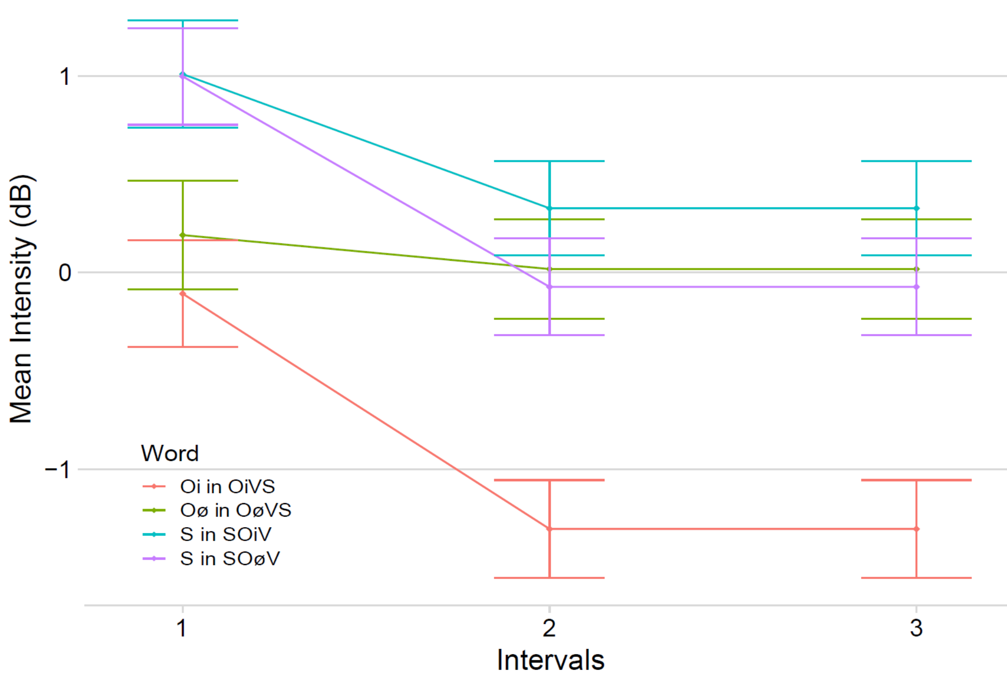


Figure 8. Normalized intensity of the first constituent in the 4 sentence types. Bars represent standard errors.

3.3. Summary

The results of this study can broadly be summarized as follows. First, there was a general tendency for Turkish sentences to decline in frequency and in intensity as they progressed. Second, there were prosodic differences between SOV and OVS sentences. Third, for both SOV and OVS sentences, there were prosodic differences between casemarked and non-casemarked versions of the sentences, with the majority of the participants exhibiting the differences. Overall, the prosodic differences were more robust for fundamental frequency than intensity, and for most participants the differences were greater for F0 than intensity. Lastly, the pitch contours of the first words of scrambled and non-scrambled sentences were noticeably different for non-casemarked sentences, whereas the pitch contours of the first words of scrambled and non-scrambled sentences were virtually identical for casemarked sentences. Taken as a whole, the results of this study suggest that Turkish speakers prosodically differentiate between sentences when morphosyntactic cues like word order and case-marking, and non-morphosyntactic cues like propositional content and discourse context are absent.

4. Discussion

In this study, native Turkish speakers read sentences in which overt case-marking and word order were manipulated, and propositional content and discourse context provided no clues about noun phrases' grammatical roles. Although the results of our study indicate that Turkish speakers said casemarked and non-casemarked sentences, and OVS and SOV sentences in prosodically distinct ways, the critical question is, *why* do they do so. What do our findings reveal about the nature and cognitive underpinnings of language and language processing?

4.1. Why do these prosodic differences exist?

4.1.1. Sentence processing, information theory (IT), and prosody

As discussed in the Introduction, a key step in interpreting the meaning of a sentence is to determine what the grammatical roles of its nouns are. Case-marking, word order, propositional content, discourse content and prosody all provide potential clues about nouns' grammatical roles (and hence, their thematic roles). Perhaps when more robust morphosyntactic cues for grammatical and thematic roles are not present, speakers provide probabilistic prosodic cues to help their listeners understand what they are saying. In other words, the differences in how our Turkish speakers said sentences that did and did not have overt case-marking and default word order could reflect the communicative interplay between speakers and listeners.

Since Turkish sentences are 2.5 times as likely to begin with subjects than objects (Batman-Ratyosyan, 2003), Turkish listeners probably are strongly biased to parse the first noun of sentences as a subject. From a sentence processing perspective, in order to prevent their listeners from going down this garden path, Turkish speakers may prosodically mark the first noun in object-initial sentences with a prominent pitch accent to signal to their listener that the noun is an object and not a subject.[8] If Ipek & Jun (2013) are correct in their account of rises of F0 as indicators of stress, then our finding of steeper rise and decline for object-initial sentences would be consistent with this being an indication of stress.

A similar explanation could account for the observed prosodic differences between the OV segments of Turkish sentences that did and did not have overt accusative case-marking. In flexible word order languages, morphological case-marking is a more reliable cue for a noun's grammatical role than word order (MacWhinney, Bates & Kliegl, 1984). In Turkish, when a noun has an overt accusative case marker, that noun is always the object of the sentence. However, because Turkish subjects do not receive overt nominative case, a bare noun (i.e., a noun that does not have overt morphological case) could be an object in a sentence that has SOV or OVS word order (i.e., an $SO_øV$ or $O_øVS$ sentence), or it could be a subject in a sentence with any of the possible six word orders. The more prominent pitch accent on objects that lack an overt accusative case marker may reflect that the Turkish speaker is consciously or unconsciously attempting to indicate the grammatical role of a noun that is otherwise ambiguous and, therefore, might lead the listener to initially misparse the sentence.

From an information theoretic (IT) perspective, these prosodic differences may reflect the informativeness of word order and overt case-marking. The informativeness of

[8] Although it might seem implausible that listeners would perceive this prosodic difference, speech perception research has shown that listeners are extremely successful at compensating (i.e., normalizing) for individual speaker variation in even very brief periods of time. (see Kleinschmidt & Jaeger, 2015; Kleinschmidt, 2019)

a particular element of a message can be measured by how unexpected it is (Shannon, 1948). Two recent instantiations of IT in the psycholinguistic domain are surprisal (see Hale, 2001; Boston, Hale, Vasishth & Kliegl, 2011) and entropy reduction (Genzel & Charniak, 2002, 2003) theories. Both of these theories view linguistic input (e.g., morphemes, words, sentences, etc.) as fundamentally a random variable whose conditional probability can be calculated given what has occured prior to that linguistic input (often referred to as context) and the entire set of possible inputs.

Although they differ in the mathematical definition of the particular complexity metrics they use, surprisal and entropy both attempt to quantify the 'amount of information' a message, or a particular piece of that message contains (Hale, 2016). Because most Turkish sentences begin with a subject, it is informationally dense when the first noun of a Turkish sentence is an object (or theme) and not a subject (or agent). Similarly, because most Turkish nouns that do not have overt morphological case are subjects and not objects (see 1.6.1. in Introduction), if a noun that lacks overt case-marking is an object and not a subject, this is valuable information. From this perspective, one can argue that the prosodic differences between casemarked and non-casemarked, and OVS and SOV versions of the same sentences are a natural consequence of the way information unfolds in real-time, with speakers adjusting the way they speak to keep the rate of information transfer constant (see Levy & Jaeger, 2007 and Frank & Jaeger, 2008 for details on Uniform Information Density). It would be reasonable to expect that speakers would prosodically accentuate the unexpected part (i.e., high in surprisal or potential entropy reduction) and downplay the expected part (i.e., low in surprisal or potential entropy reduction) of a message to maximize the efficiency of communication: when other methods of informing the listener about who is doing what to whom are not available (i.e., there is no overt case-marking or discourse context and the propositional content of a sentence is neutral), the speakers in our study made that distinction through prosody.

The IT perspective is, however, agnostic as to the origins of the prosodic differences. It could be that the speaker consciously knows and has planned ahead, which part of the sentence is most important and therefore knowingly emphasizes the part which is the most informative (i.e., has the highest surprisal or entropy reduction potential). On the other hand, it could be that over time, communications systems evolve to maximise the efficiency, either in terms of decreasing surprisal/entropy rapidly or retaining a uniform rate of information transfer, or both, in which case the speaker would be unaware that they produce overtly casemarked and non-casemarked sentences and OVS and SOV sentences in prosodically differently ways.

4.1.2. Sentence production and prosody

Perhaps the prosodic differences we found reflect the processes involved in sentence production. Previous research on sentence production (e.g., Konopka, 2012) suggests that the scope of planning for a phrase is much smaller for sentences that start with a lower-frequency word compared to sentences that start with a higher frequency word. Konopka defines planning scope as the amount of linguistic information one prepares to utter, and argues that "the ease of lexical encoding and *structural* formulation" (p. 152, emphasis added) influences the production onset and consequently the planning scope of chunks of utterances. The assumption is that this finding reflects that producing less common words and *structures* involve greater or different resources, mechanisms, or routes than producing more common ones.

According to this perspective, the fact that the SOV word order is more common than the OVS word order would mean that the planning scope of SOV sentences is larger than the planning scope of OVS sentences, and this could account for SOV and OVS sentences being prosodically different. Similarly, according to this account, because casemarked sentences are more common in Turkish than non-casemarked sentences, the planning scope for casemarked sentences should be greater which might result in them being produced in a prosodically distinct way from non-casemarked sentences. However, it is unclear why the differences in planning scope would result in the specific prosodic contours we observed. Why, for example, should the smaller planning scope cause the non-casemarked objects to rise higher and decline faster than casemarked objects? Similarly, why is the peak in the first noun in OVS sentences more pronounced than that in SOV sentences? Teasing apart whether the the differences are due to the effects of frequency and thus planning scope, or due to the syntactic structure of sentences is confounded by the fact that the structures that we would like to compare occur at different frequencies in natural languages (e.g., SOV is more common than OVS). This renders designing experimental sentences that separate the effects of frequency and structure a rather challenging if not unsolvable problem.

4.1.3. Syntax and prosody

Could the observed prosodic differences -- between sentences that do and do not have default word order and between sentences that do and do not have overt accusative case-marking -- be linguistic in origin rather than psycholinguistic/information theoretic? Put another way, could they (at least partially) reflect the structure of language rather than merely reflecting the processes involved in producing and processing language?

For example, perhaps the more prominent pitch accent on the first word of sentences when the first word is an object rather than a subject reflects differences in the tree structures of SOV and OVS sentences. If, as many linguists have argued, the underlying word order of all Turkish sentences is SOV, this means that in OVS sentences, subject NPs have "moved" to a post-verbal position. Thus, although an object in an OVS sentence and a subject in an SOV sentence are both the first noun in their respective sentences, they occupy different positions within their respective syntactic trees. In SOV sentences, the subject is in the first DP node of the sentence, whereas in OVS sentences, the object is the second NP and there is an empty DP node before the object where the subject formerly was[9].

Let us now turn to the observed prosodic differences between casemarked and non-casemarked objects. Some grammatical theories argue that all nouns must receive case (Chomsky, 1982). Perhaps in addition to case being morphologically realized with a suffix, case can be prosodically realized. Alternatively, it could be that when an object does not have an overt morphological casemarker, it must remain strictly within the domain of the verb, whereas when an object has an overt accusative casemarker, it is free(er) to move to different nodes. If this is the case, then the differences between overtly and non-overtly casemarked sentences could reflect the interaction between prosody and syntax.

As discussed in section 1.6.2. of the Introduction, SOV sentences that do and do not have overt case-marking both become disambiguated at the end of the second noun. In

---

[9] There is disagreement about the exact configuration of scrambled sentences (Kornfilt 2003; 2013). Some researchers argue that it is also possible that, in OVS sentences, the OV has moved to a pre-subject position. For the purposes of this argument, what matters is that the structures of SOV and OVS sentences differ, not the exact way in which they differ.

striking contrast, OVS sentences that have overt case-marking (i.e., $O_{ACC}VS$) become disambiguated after the first noun, whereas OVS sentences without overt case-marking (i.e., $O_øVS$) do not become disambiguated until the very end of the sentence. Given this, if speakers are consciously or unconsciously providing clues to help their listener avoid garden paths, we would predict that the prosodic marking of non-casemarked objects would be more prominent in $O_øVS$ than $SO_øV$ sentences. The fact that we observed the same prosodic differences between overtly casemarked and non-casemarked sentences for both SOV and OVS sentences suggests that these differences associated with non-overt casemarking may not merely be helpful tips given by the speaker, but may be partially linguistic in nature.

4.2. Future experiments

Based on the production data we collected, it is not possible to determine whether the prosodic differences we observed among the four sentence types reflect the processes involved in producing or processing these sentences, the informativeness and information density of these sentences, and/or the structure of these sentences. In the final section of this thesis, we will outline some experiments that could potentially shed light on the cause(s) of our prosodic findings.

Comprehension studies. As mentioned previously, the online interpretation of casemarked and non-casemarked OVS sentences differ in that casemarked $O_{ACC}VS$ sentences are disambiguated after the first noun while non-casemarked $O_øVS$ sentences are not disambiguated until the end of the last word. From a processing/IT perspective, the early prosodic differentiation between casemarked and non-casemarked NVN sentences might be an indication that speakers are providing an early probabilistic prosodic cue to help listeners avoid potential garden-path effects in sequences such as $N_{ø}VN_{ø}$ in which more robust morphosyntactic cues are lacking.

We could explore whether the observed prosodic differences are clues to help the listener avoid potential garden-path sentences by conducting comprehension studies to determine whether the presence of these cues actually help Turkish speakers understand sentences that are scrambled or lack case-marking. In our study, some speakers (for example, participant 1) provided more robust prosodic cues than others (for example, participant 9). We plan to harness this variability by investigating whether native Turkish-speaking adults' comprehension of scrambled and non-casemarked sentences are affected by the robustness of individual speaker's prosodic cues. If listeners recognize and rely on the prosodic differences to parse sentences and avoid garden paths, then we would expect listeners to be faster and more accurate at understanding sentences said by Turkish speakers who provide more robust prosodic cues. If listeners do not rely on these prosodic differences to parse the sentences, then we would expect them to have similar reaction times and accuracy rates regardless of the robustness of speakers' prosodic differences.

Another possible technique would be to take a speaker's sentences and artificially manipulate them such that the prosodic differences are accentuated in some sentences and not in others. If Turkish-speaking adults use prosody to avoid garden paths, we would expect them to perform better when they listen to the sentences in which the prosodic cues are exaggerated. Finally, one can artificially manipulate the prosodic profile of sentences to be *incongruent* with their syntactic structure. For example, the object in a non-casemarked declarative OVS sentence could be modified to have the pitch contour associated with subjects in subject-initial sentences or the pitch contour associated with objects in casemarked OVS sentences. If people use prosodic cues to guide their parsing, we would expect them to be slower and less accurate when the prosodic cues were incongruent with the structure.

We could also conduct eye-tracking studies to see whether there are differences in the eye-gaze patterns of listeners who are tasked with identifying the agent and the theme in a sentence they hear (Özge, Küntay & Snedeker, 2019). For example, in a visual world paradigm we could show listeners two pictures that describe the same action, with opposing thematic roles (e.g., a man serving a woman, vs. a woman serving a man).

If people rely on early prosodic cues to correctly parse potential garden-path NVNs (i.e., $O_{ø}$VS sentences), then we would expect more early shifts in listeners' gaze to the correct picture after hearing the first noun when these NVNs are uttered by speakers who display robust prosodic cues for object case-marking compared to when the sentences are uttered by speakers who do not provide robust prosodic cues. If listeners do not rely on prosodic cues or if prosodic cues are absent, then they might go down the garden path and interpret the sequence to be an $SVO_{ACC,}$ sentence, in which case, they will be forced to reanalyze $O_{ø}$S sentences upon encountering the final noun which also lacks case-marking. In this case we would expect more early looks to the incorrect picture, and possible multiple saccades between the two pictures caused by reanalysis. Alternatively, listeners could wait until the end of the sentence to process its meaning, in which case we would expect equal proportions of looks to correct and incorrect pictures throughout the course of the sentence.

Production Studies. If the prosodic differences we observe between case-marked and non-casemarked sentences and SOV and OVS sentences are motivated by speakers consciously or unconsciously providing cues for their listeners, then we would expect that speakers would amplify the prosodic differences in experiments that accentuate the communicative nature of the task. If, on the other hand, our prosodic findings simply reflect differences in the structures of the four sentence types, we would expect speakers to say the sentences the same way regardless of the experimental task demands.

The easiest way to investigate this simply involves manipulating the instructions given to speakers, to determine whether doing so influences how speakers produce sentences. For example, half of the participants could be told that they are being recorded to test out our audio equipment, and the other half could be told that they are making recordings that will be used in future comprehension experiments. If psycholinguistics/IT factors are the cause of these prosodic differences, we would predict that there would be greater prosodic differences between casemarked and non-casemarked sentences and between SOV and OVS sentences for speakers who are told they are making recordings that will be used in future comprehension studies than for speakers who are told they are reading sentences aloud to test our audio equipment. If, on the other hand, the observed prosodic differences are simply due to structural differences among the four sentence types, we would not expect any differences between the prosodic contours of the sentences for the two groups of speakers.

Another way to test the extent to which the observed differences are due to psycholinguistic versus structural factors is to design an experiment in which speakers either read sentences to another person or read the sentences aloud to themselves. We could, for example, bring in pairs of native Turkish speakers and have them engage in a "game" in which the amount the speaker-listener pair gets paid depends on how many sentences the listener correctly understands in a set amount of time. This would incentivize speakers to provide prosodic cues that help their listeners avoid going down potential garden paths. Psycholinguistics/IT account would predict that speakers will provide more robust prosodic cues when they are engaged in this sort of conversational "game" than when they are simply reading sentences aloud to themselves. In contrast, if the prosodic differences are primarily linguistic in origin, we would not expect to see any differences in the prosody of the sentences in the two experimental conditions.

Research suggests that speakers may provide visual cues that align (or sometimes conflict with) with auditory cues (Bosker & Peeters, 2020). For example, facial mimicry (Garg, Hamarneh, Jongman, Sereno & Wang, 2019), body language, hand gestures and sustained gaze (i.e., eye contact) appear to co-occur with auditory/linguistic cues in a way that signals the 'crux' of the message. If speakers use whatever channels are available to communicate clearly and efficiently with each other, then when some channels are not available (e.g., telephone calls where visual cues are absent or text messages where auditory cues are absent), we would expect them to accentuate cues or use additional cues in the available channels (e.g., prosody, emojis).

We could take advantage of this by manipulating the channels of information that are available and studying whether the strength of prosodic cues differs depending on what channels are available. For example, in a study in which participants are told they are reading sentences that will be used in future comprehension studies, we could compare the strength of the prosodic differences when speakers are only audio-recorded versus when they are simultaneously audio- and video-recorded. If visual cues like body language, gestures or facial mimicry help people process sentences and avoid garden paths, then psycholinguistic/IT accounts might predict that speakers will provide stronger prosodic cues in the audio-only condition because visual cues are not available. On the other hand, if our prosodic findings simply reflect structural differences among the four sentence types, the prosodic differences among the sentences should be similar in magnitude in the only audio and audio plus video conditions.

One could also argue that psycholinguistic/IT accounts would predict that the quality of the audio channel might affect the strength of the prosodic cues. For example, the psycholinguistics/IT account would predict that speakers in noisy environments might accentuate the prosodic differences among the four sentence types compared to speakers in quieter environments. If, on the other hand, the differences among the sentence types simply reflect structural differences among the sentence types, the noisiness of the recording environment should not have a consistent effect on the prosody contours of sentences.

As discussed in section 1.4.2, contextual information could potentially help listeners avoid garden-paths. Therefore, we could explore whether the observed prosodic differences reflect psycholinguistic/IT factors or linguistic factors by providing contextual information that reduces the likelihood of listeners going down garden paths. For example, the contextual information in Table 6 makes the agent and the theme of the target sentence (i.e., underlined sentence) perfectly predictable in all four sentence types. According to the psycholinguistic/IT perspective, contextual information preceding the target sentence should reduce the overall entropy and surprisal previously associated with the first word, because the preceding contextual information clearly determines the thematic roles in the target sentence.

$SO_{ACC}V$	Parkta bir oğlan ve bir kız kovalamaca oynuyorlardı. Önce kız oğlanı kovaladı. Sonra da rolleri değiştiler. Oğlan kızı kovaladı.
	In the park, a boy and a girl were playing catch. First the girl chased the boy. Then they switched roles. The boy chased the girl.
$SO_{\emptyset}V$	Babası Cenk'in geçmişteki davranışlarından memnun olmadığını söyledi: "Cenk okuldaki derslerine dikkat etmesi gerektiği yerde vaktini boşa harcadı. Oğlan kız kovaladı."
	His father said he was not satisfied with Cenk's behaviour in the past: "Cenk squandered his time when he should have been paying attention to his classes. The boy chased (after) girl(s)."
$O_{ACC}VS$	Parkta bir oğlan ve bir kız kovalamaca oynuyorlardı. Önce kız oğlanı kovaladı. Sonra da rolleri değiştiler. Kızı kovaladı oğlan.
	In the park, a boy and a girl were playing catch. First the girl chased the boy. Then they switched roles. It was the girl that the boy chased.
$O_{\emptyset}VS$	Babası Cenk'in geçmişteki davranışlarından memnun olmadığını söyledi:"Cenk okuldaki derslerine dikkat etmesi gerektiği yerde vaktini hep karşı cinsle harcadı. Kız kovaladı oğlan."
	His father said he was not satisfied with Cenk's behaviour in the past: "Cenk squandered his time with the opposite gender when he should have been paying attention to his classes. It was girl(s) that the boy chased (after)."

Table 6. Sentences with preceding context. The surprisal and the possible entropy reduction at target sentences are low given the obvious context.

If the prosodic differences among the four sentence types are smaller when contextual information is provided, then this would indicate that these differences are part of the speakers toolkit to facilitate smoother communication and help listeners avoid garden paths. If the prosodic differences among the four sentence types merely reflect structural differences among them, then embedding sentences in different contexts should not affect how the speakers utter them.

Recall from section 1.4.1 that the propositional content of sentences can aid people in parsing them. Because this study was designed to investigate how morphosyntactic factors affect prosody, we controlled for propositional content by having participants read sentences that always had two animate nouns for subject and object, either of which could plausibly be the agent of the action described by the verb. We could conduct a production experiment in which some of the sentences that the participants read have semantically reversible arguments (e.g., *The child saw the mother*) and some that do not (e.g., *The child saw the accident)* to gain insight into the cause of the prosodic differences. The psycholinguistics/IT account predicts that if the thematic roles of the arguments can be recovered merely by knowing the meanings of the nouns and verbs (e.g., chew, bone, dog), speakers should provide less robust prosodic cues compared to sentences where the morphosyntactic information is required to determine the thematic roles of the arguments (e.g., child, mother, see). The linguistic account again predicts that the prosodic differences we observed for the four types of sentences should be similar in strength for semantically reversible and irreversible sentences because the syntactic structure of the four sentence types is independent of the meanings of the nouns and whether they can plausibly switch thematic roles (i.e., child-mother type pairs) or not (i.e., child-accident type pairs).

Recall that we observed the same prosodic difference between casemarked and non-casemarked objects in SOV and OVS sentences even though SOV sentences are disambiguated at the same point in casemarked and non-casemarked sentences. This suggests that the prosodic differences we observed may (at least partially) reflect the linguistic structure of sentences, rather than simply being acoustic “breadcrumbs” left by the speaker to guide the listener down the correct path. We could further test whether the observed prosodic differences among the four types of sentences are partly due to the structure of these sentences by having speakers produce other types of sentences. For example, if prosody is an alternative way of marking an object’s case, then SVO, OSV, VOS, and VSO sentences that do not have overt morphological accusative case should be acceptable if the object is pronounced with a prominent pitch accent. One way to test this would be to have native Turkish speakers judge the acceptability of ungrammatical SVO$_{ø}$, O$_{ø}$SV, VO$_{ø}$S, and VSO$_{ø}$ sentences when the non-casemarked objects in these sentences do and do not have the prosodic characteristics we found in non-casemarked SOV and OVS sentences. If native speakers give high acceptability ratings to SVO$_{ø}$, O$_{ø}$SV, VO$_{ø}$S, and VSO$_{ø}$ sentences produced with a prominent pitch accent on the object, then this could be considered evidence towards the linguistic account in which prosody is an alternative way of marking the case of an object.

Most studies on sentence processing have investigated how morphosyntax and semantic factors such as animacy affect sentence processing. These studies have also tended to be conducted on languages like English which has strict word order. The study presented in this thesis attempts to fill these gaps in the literature by studying the prosodic features of Turkish, a language with flexible word order and rich morphology. In addition, this study and the proposed studies provide a promising way of exploring how sentences differ prosodically, and the extent to which sentential prosody reflects the mechanisms involved in producing or processing sentences, the informativeness and information density of sentences, and their structure.

Bibliography

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*(3), 191-238.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*(3), 355-387.

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, **67**(1), 1–48. doi: 10.18637/jss.v067.i01.

Batman-Ratyosyan, N. (2003). The acquisition of word order and morphology in Turkish. Ph.D. Dissertation. Rutgers University.

Batmanian, N. & Stromswold, K. (2019) Word order rules: Parsing sentences in "free" word order languages. *Proceedings of the 43th Annual University of Pennsylvania Linguistics Conference*, volume 26.1.

Bever, T. G. (1970). The cognitive basis for linguistic structures. *Cognition and the Development of Language*, *279*(362), 1-61.

Birch, S., & Clifton, C. (2002). Effects of varying focus and accenting of adjuncts on the comprehension of utterances. *Journal of Memory and Language*, *47*(4), 571-588.

Blake, B. J. (2001). Case. 2nd edition. Cambridge: Cambridge University Press

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic Bulletin & Review*, *18*(6), 1189-1196.

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer [Computer Program]. Version 6.0. 39 retrieved in May 2018.

Bosker, H. R., & Peeters, D. (2020). Beat gestures influence which speech sounds you hear. *bioRxiv*.

Boston, M. F., Hale, J. T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, *26*(3), 301-349.

Breen, M. (2014). Empirical investigations of the role of implicit prosody in sentence processing. *Language and Linguistics Compass*, *8*(2), 37-50.

Carlson, K. (2009). How prosody influences sentence comprehension. *Language and Linguistics Compass*, *3*(5), 1188-1200.

Chomsky, N. (1964). Current Issues in Linguistic Theory (The Hague: Mouton). *Pragmatism and Binding/279*.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. MIT press.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*(4), 368-407.

Clifton, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In *Linguistic Structure in Language Processing* (pp. 273-317). Springer, Dordrecht.

De, C. A., Bolinger, D. L., Gibbon, D., Garding, E., T'Hart, J., Gronnum, N., Alcoba, S., & Murillo, J. (1998). *Intonation Systems: a Survey of Twenty Languages*. Cambridge University Press.

Dryer, M. S. (2013) Order of Subject, Object and Verb. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) The World Atlas of Language Structures Online. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/81, accessed on 2019-11-23

Erguvanli, E. E., & Taylan, E. E. (1984). *The function of word order in Turkish grammar* (Vol. 106). Univ of California Press.

Erguvanli, E., & Zimmer, K. (1994). Case marking in Turkish indefinite object constructions. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 20, No. 1, pp. 547-552).

Fillmore, C. J. (1968). Lexical Entries for Verbs. *Foundations of Language*, 373-393.

Fodor, J. D. (1989). Empty categories in sentence processing. *Language and Cognitive Processes*, *4*(3-4), SI155-SI209.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In *Speech Prosody 2002, International Conference*.

Frank, A. F., & Jaeger, T. F. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 30, No. 30).

Frazier, L. (1978) On comprehending sentences: syntactic parsing strategies. *Doctoral dissertation, University of Connecticut*.

Frota, S. (2002). The prosody of focus: a case-study with cross-linguistic implications. In *Speech Prosody 2002, International Conference*.

Garg, S., Hamarneh, G., Jongman, A., Sereno, J. A., & Wang, Y. (2019). Computer-vision analysis reveals facial movements made during Mandarin tone production align with pitch trajectories. *Speech Communication*, *113*, 47-62.

Genzel, D., & Charniak, E. (2002). Entropy rate constancy in text. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics* (pp. 199-206).

Genzel, D., & Charniak, E. (2003). Variation of entropy and parse trees of sentences as a function of the sentence number. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing* (pp. 65-72).

Godfrey, John J., and Edward Holliman. Switchboard-1 Release 2 LDC97S62. Web Download. Philadelphia: Linguistic Data Consortium, 1993.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. *In J. Greenberg, ed., Universals of Language. 73-113. Cambridge, MA.*

Hahn, M., Jurafsky, D., & Futrell, R. (2020). Universals of word order reflect optimization of grammars for efficient communication. *Proceedings of the National Academy of Sciences*, *117*(5), 2347-2353.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Second Meeting of the North American Chapter of the Association for Computational Linguistics*.

Hale, J. (2016). Information-theoretical complexity metrics. *Language and Linguistics Compass*, *10*(9), 397-412.

Ipek, C., & Jun, S. A. (2013). Towards a model of intonational phonology of Turkish: Neutral intonation. In *Proceedings of Meetings on Acoustics ICA2013* (Vol. 19, No. 1, p. 060230). Acoustical Society of America.

İşsever, S. (2003). Information structure in Turkish: the word order–prosody interface. *Lingua*, *113*(11), 1025-1053.

Kabak, B., & Vogel, I. (2001). The phonological word and stress assignment in Turkish. *Phonology*, *18*(3), 315-360.

Kabak, B., Maniwa, K., & Kazanina, N. (2010). Listeners use vowel harmony and word-final stress to spot nonsense words: A study of Turkish and French. *Laboratory Phonology*, *1*(1), 207-224.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass*, *2*(4), 647-670.

Kan, S. (2009). Prosodic domains and the syntax-prosody mapping in Turkish. *Master's thesis. Boğaziçi University*.

Kerstens, J. (2012). *The syntax of number, person and gender: A theory of phi-features* (Vol. 18). Walter de Gruyter.

Kleinschmidt, D. F. (2019). Structure in talker variability: How much is there and how much can it help?. *Language, Cognition and Neuroscience*, *34*(1), 43-68.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148.

Knutsen, S., Stromswold, K., & Kleinschmidt, D. (2019). Segmental duration as a cue to sentence structure. *The Journal of the Acoustical Society of America*, *145*(3), 1910-1910.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, *66*(1), 143-162.

Kornfilt, J. (2003). Scrambling, subscrambling, and case in Turkish. Chapter in *Word Order and Scrambling* (Vol. 4). John Wiley & Sons, *125155*.

Kornfilt, J. (2013). *Turkish*. Routledge.

Kural, M. (1992). Properties of scrambling in Turkish. *Ms., UCLA*.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23(2), 127-150.

Nariyama, S. (2009). 'Agent Disguising Constructions from a Cross-Linguistic Perspective in Natural Language Processing. *Linguistics in the Twenty First Century*, 79.

Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Scholkopf, ¨ J. Platt, & T. Hoffman (Eds.), (p. 849-856). Cambridge, MA: MIT Press.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, *41*(5), 791-824.

Özçelik, Ö., & Nagai, M. (2011). Multiple subject positions: A case of perfect match between syntax and prosody. In *Proceedings of the 28th West Coast Conference for Formal Linguistics (WCCFL 28)* (pp. 303-312).

Özge, D., Küntay, A., & Snedeker, J. (2019). Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition*, *183*, 152-180.

Perlmutter, D. M., & Postal, P. M. (1984). Impersonal passives and some relational laws. *Studies in relational grammar*, *2*, 126-170.

Pike, K. L. (1945). The Intonation of American English.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*(3), 348-379.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002) E-Prime User's Guide. Pittsburgh: Psychology Software Tools Inc.

Sezer, E. (1983). On non-final stress in Turkish. Journal of Turkish Studies 5. 61-69

Shannon, C.E. (1948), "A Mathematical Theory of Communication", Bell System Technical Journal, 27, pp. 379–423 & 623–656, July & October, 1948.

Shintani, M. (1979). The Frequency and Usage of the English Passive, unpublished Ph. D. *Dissertation in Applied Linguistics, UCLA*.

Stromswold, K. (2006). Why children understand and misunderstand sentences: An eye-tracking study of passive sentences. *Rutgers Center for Cognitive Science Technical Report*.

Traxler, M. J., & Tooley, K. M. (2007). Lexical mediation and context effects in sentence processing. *Brain Research*, *1146*, 59-74.

Team, R. C. (2013). R development core team. *RA Lang Environ Stat Comput*, *55*, 275-286.

Turk, A. E., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics*, *35*(4), 445-472.

Vaissière, J. (1995). Phonetic explanations for cross-linguistic prosodic similarities. *Phonetica*, *52*(3), 123-130.

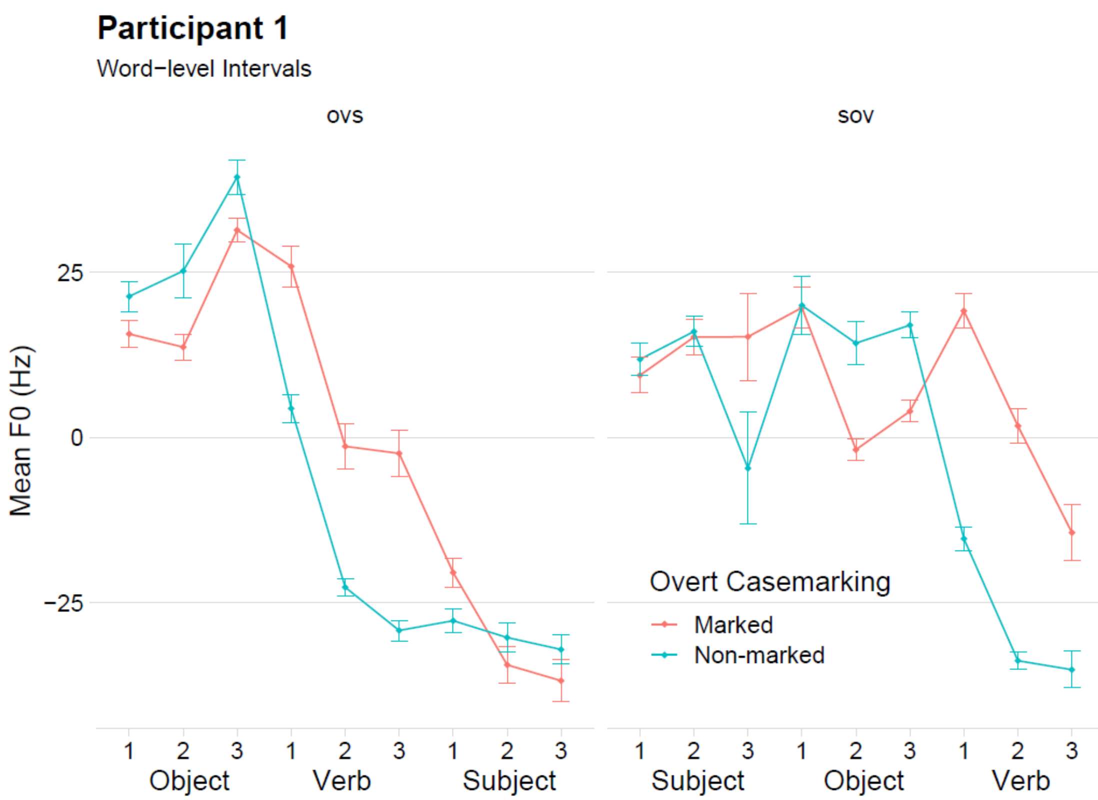
Appendix A: Experimental Scenarios Used With All Participants

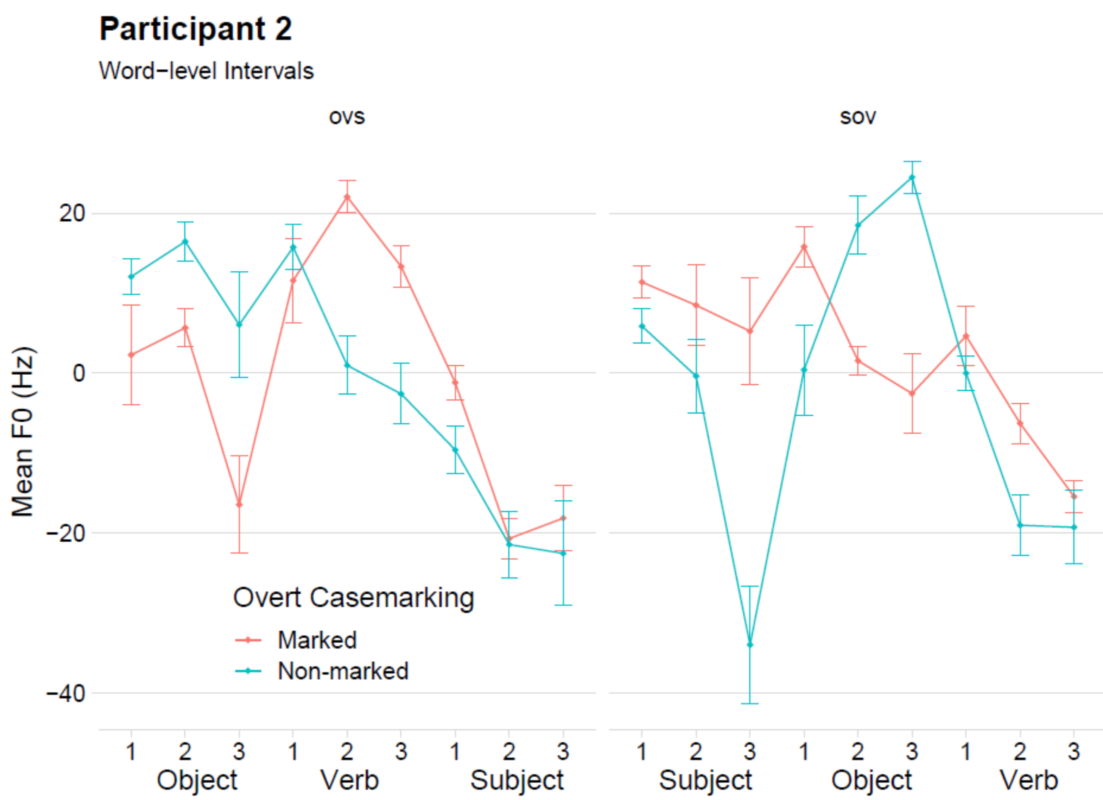
Index	**Scenarios seen by all**
1	nurse-bring-doctor
2	athlete-belittle-businessman
3	tiger-catch-monkey
4	kid-search-grownup
5	donkey-chase-ox
6	simitseller-trick-kioskowner
7	sheep-push-lamb
8	novelist-criticize-poet
9	woman-cheat.on-man
10	gardener-surprise-sweeper
11	zebra-notice-lion
12	inspector-scan-commander
13	princess-kiss-prince
14	police-take.away-chief
15	young.man-impress-woman
16	giraffe-annoy-elephant
17	editor-point.to-journalist
18	butler-spy.on-cleaner
19	merchant-stab-thief
20	man-meet-woman
21	cat-see-mouse
22	taxi.driver-kill-bus.driver
23	girl-like-boy
24	rabbit-smell-squirrel
25	student-protect-teacher
26	electrician-praise-plumber
27	minister-save-president
28	model-divorce-musician
29	fisherman-reprimand-butcher
30	deer-scare-forester

Appendix B: Experimental Scenarios Used With Only Some Participants

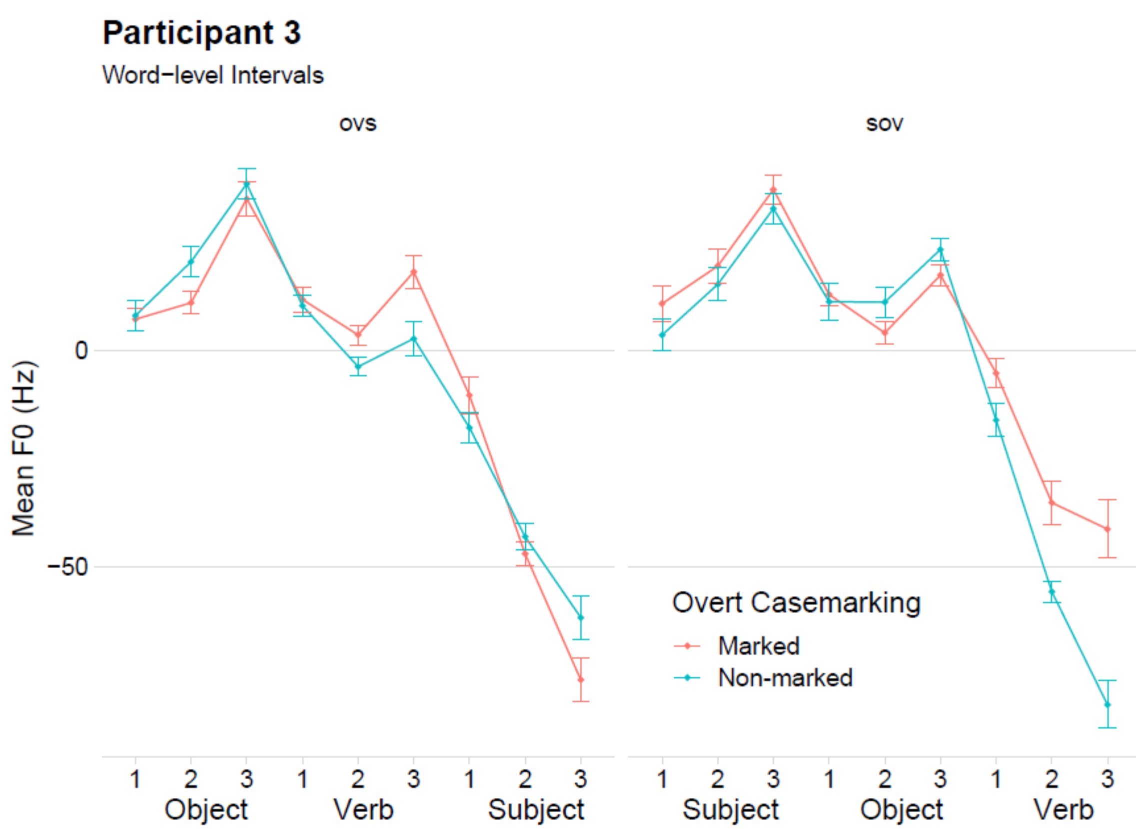
Index	**Scenarios seen by P1,P2 only**	**Index**	**Scenarios seen by all except for P1,P2**
37	uncle-spoil-aunt	31	man-watch-girl
38	elder.sister-gossip-sibling	32	boy-chase-girl
39	grandmother-shame-neighbor	33	girl-see-man
40	grandfather-watch.over-kid	34	boy-listen-woman
41	uncle-make.happy-nephew	35	woman-make.happy-kid
42	dad-embarass-son	36	villager-gossip-worker

Appendix C: Individual Participants' F0 values for OVS sentences and SOV sentences.

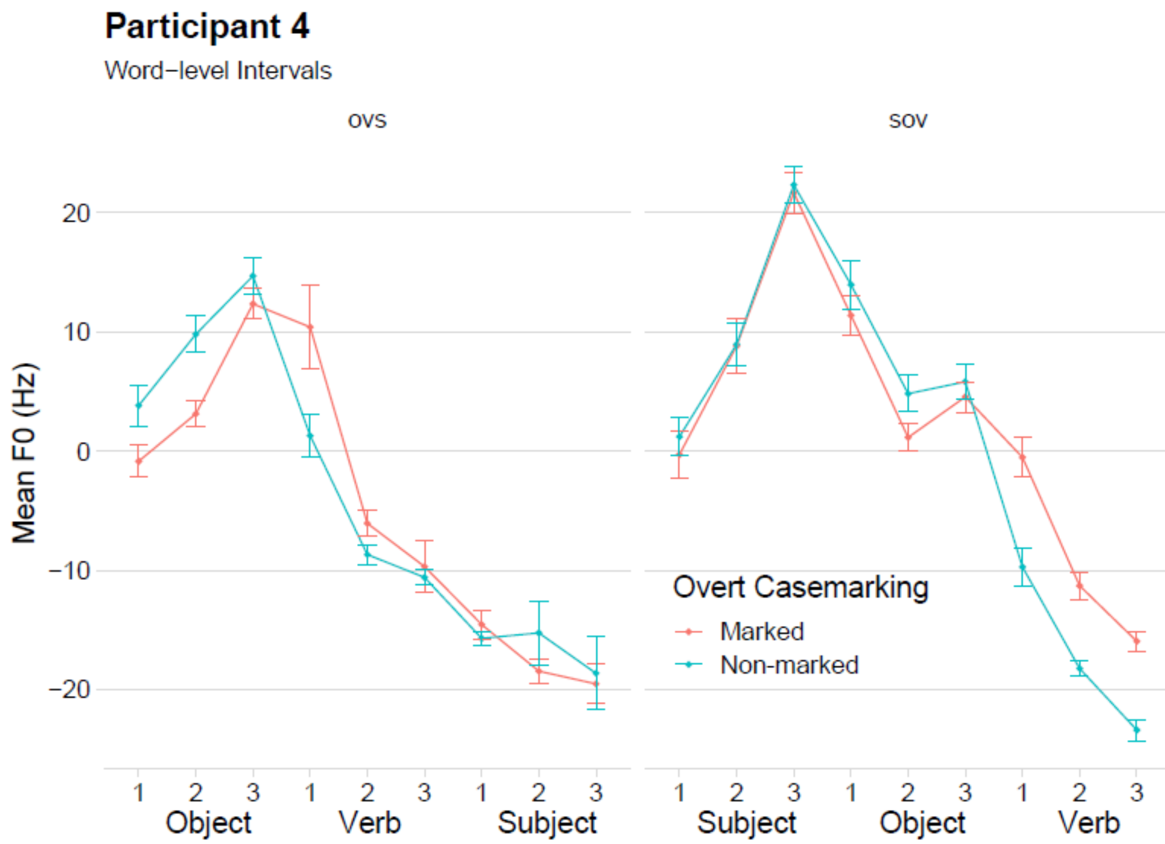


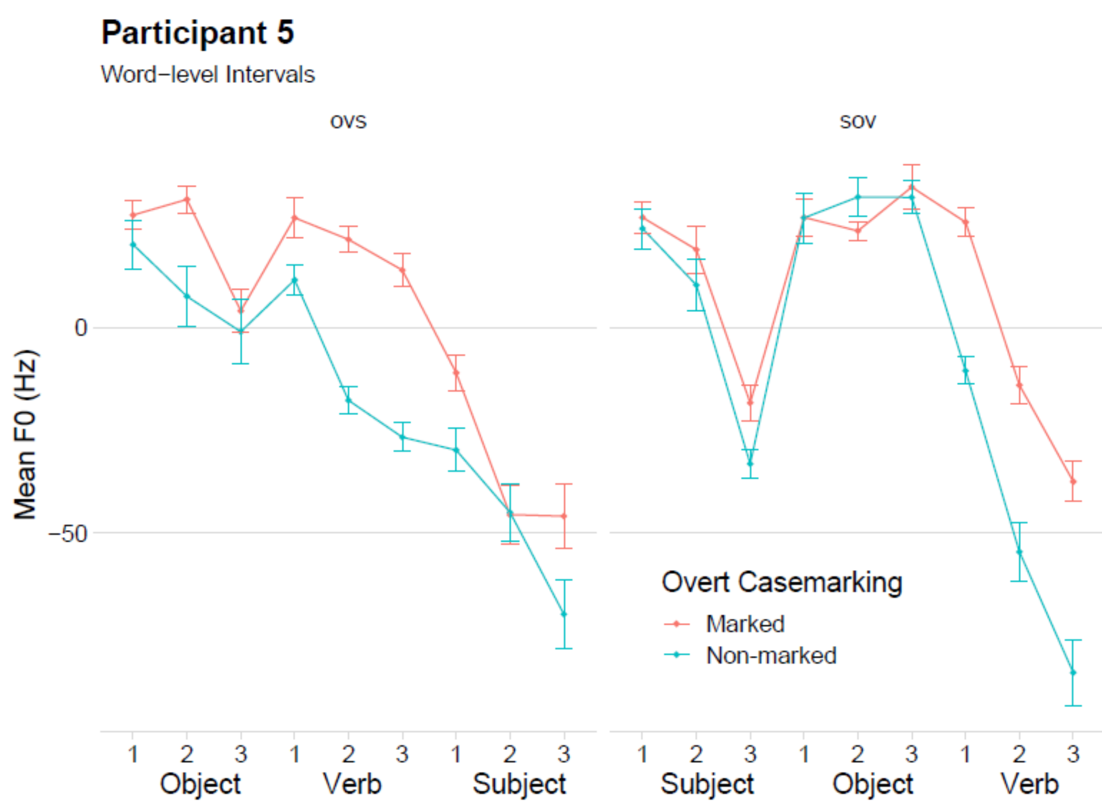


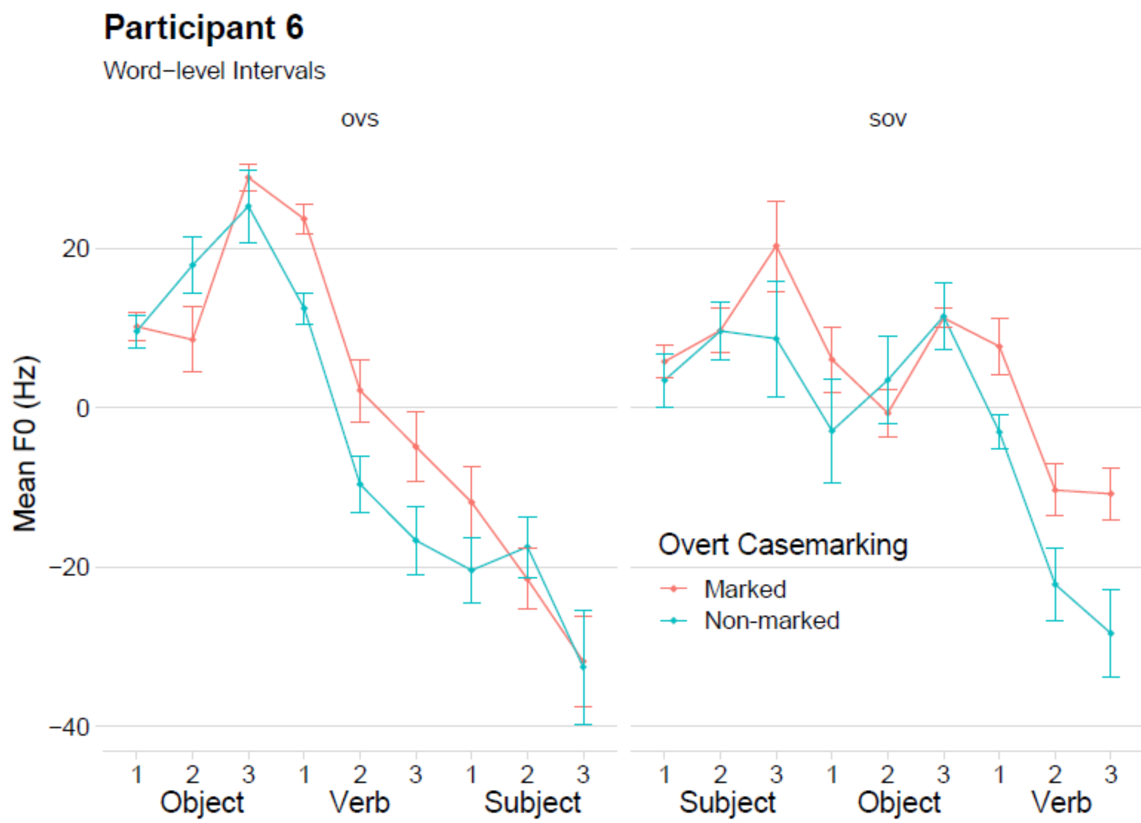
Participant 2
Word−level Intervals
ovs
sov
Mean F0 (Hz)
20
0
−20
−40
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
1 2 3 1 2 3 1 2 3
Subject
Object
Verb

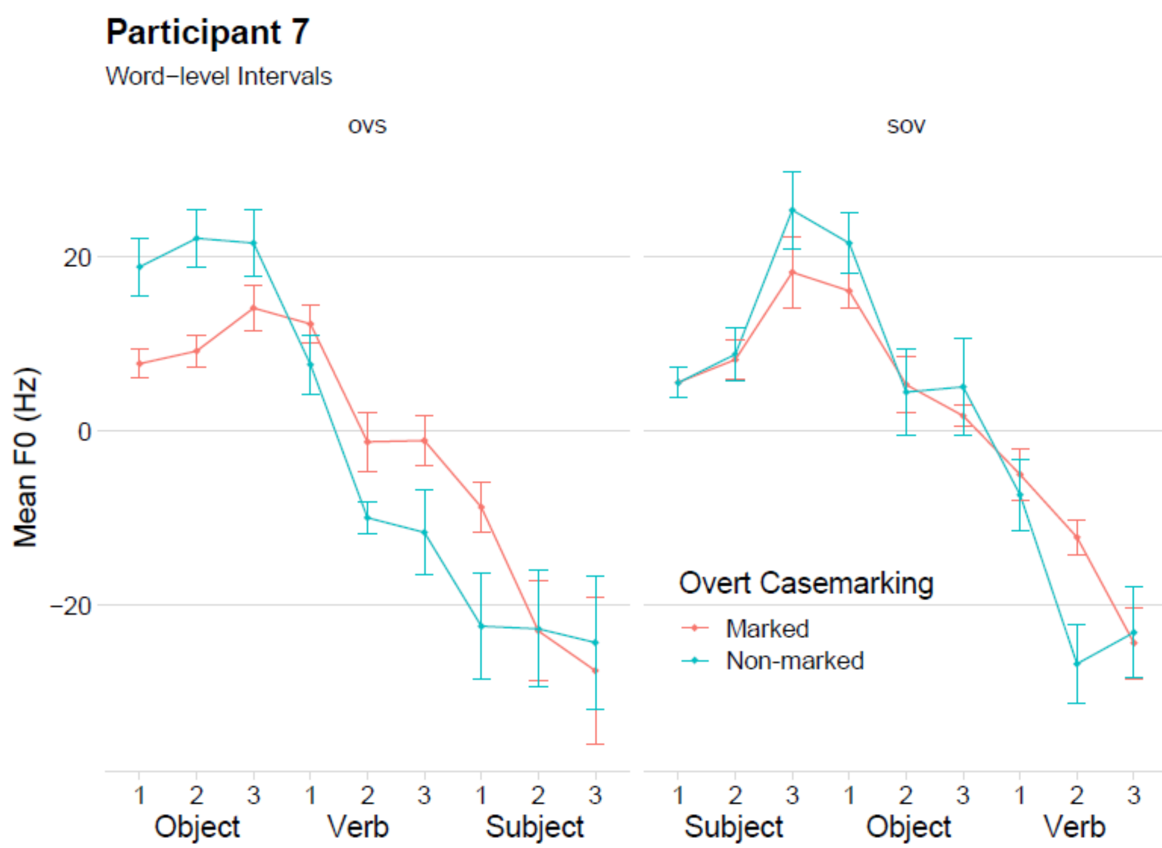


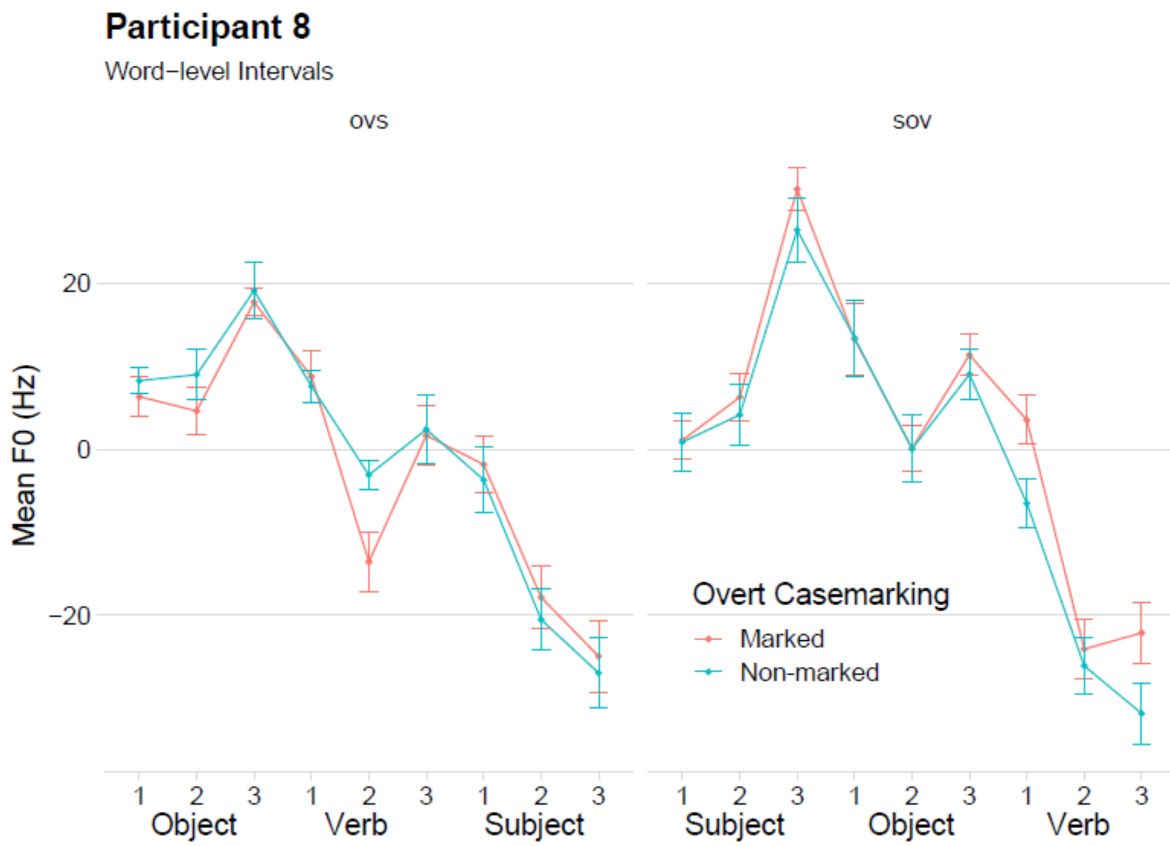
Participant 3
Word−level Intervals
ovs
sov
Mean F0 (Hz)
0
−50
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
1 2 3 1 2 3 1 2 3
Subject
Object
Verb

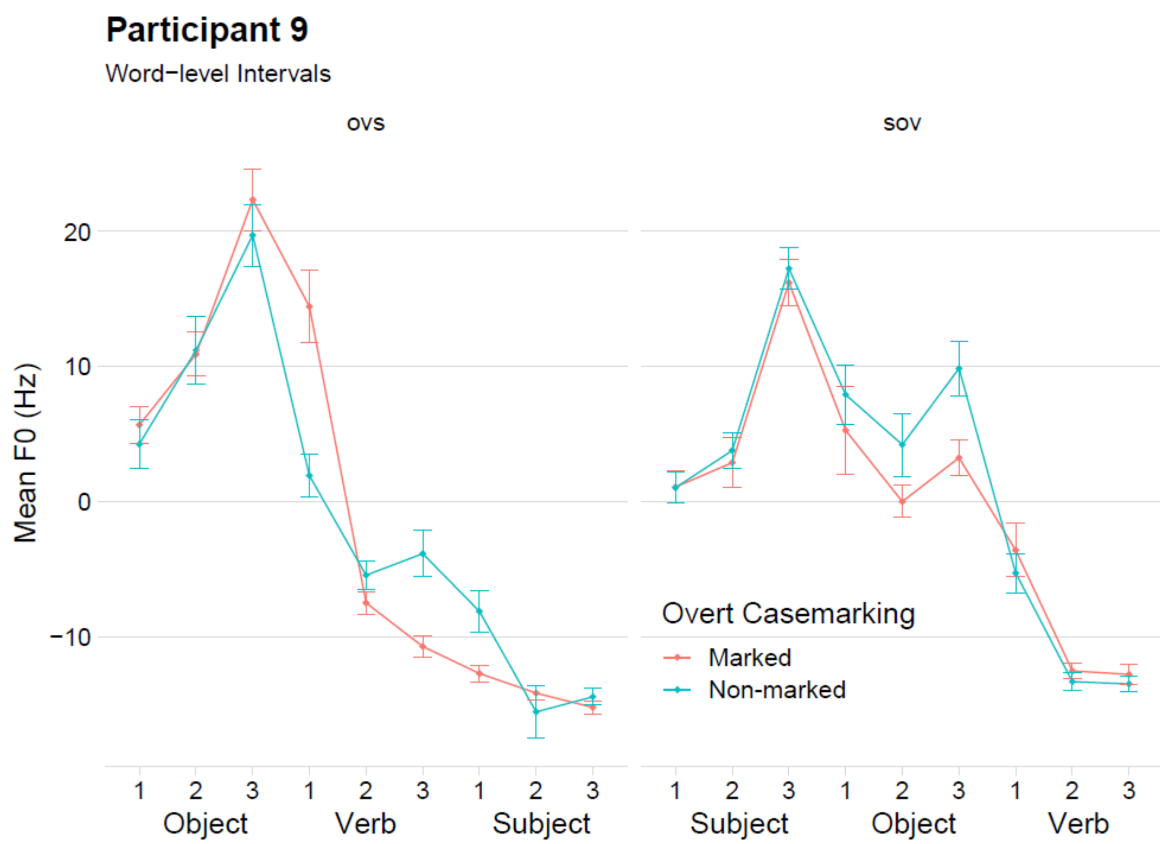

Participant 4
Word−level Intervals
ovs
sov
Mean F0 (Hz)
20
10
0
−10
−20
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object Verb Subject
1 2 3 1 2 3 1 2 3
Subject Object Verb


Participant 5
Word−level Intervals
ovs
sov
Mean F0 (Hz)
0
−50
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object Verb Subject
1 2 3 1 2 3 1 2 3
Subject Object Verb

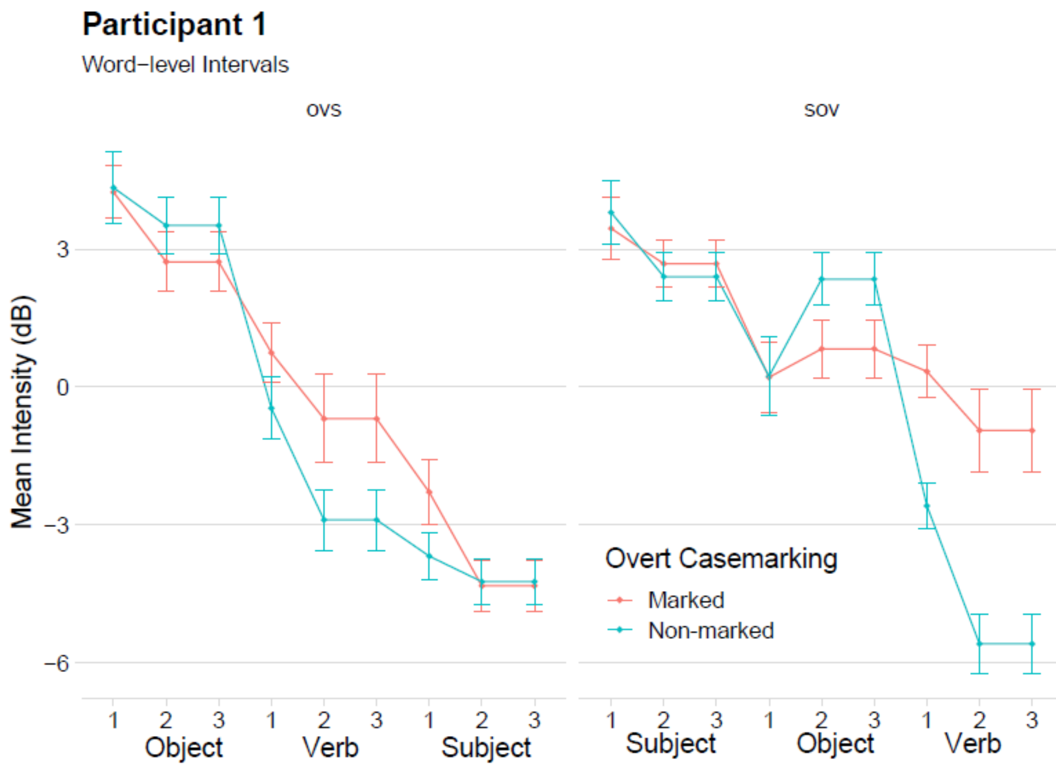

Participant 6
Word−level Intervals
OVS
SOV
Mean F0 (Hz)
20
0
−20
−40
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
1 2 3 1 2 3 1 2 3
Subject
Object
Verb
Overt Casemarking
Marked
Non-marked

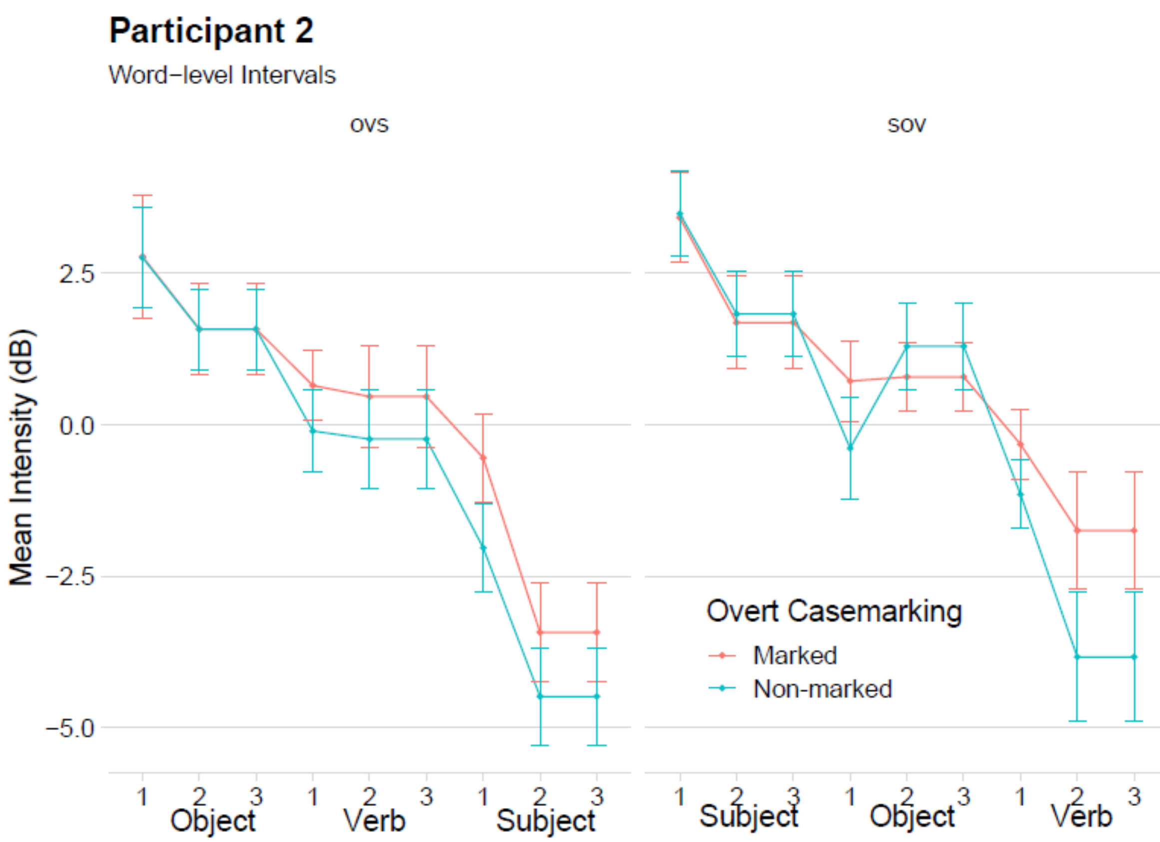

Participant 7
Word−level Intervals
OVS
SOV
Mean F0 (Hz)
20
0
−20
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
1 2 3 1 2 3 1 2 3
Subject
Object
Verb
Overt Casemarking
Marked
Non-marked

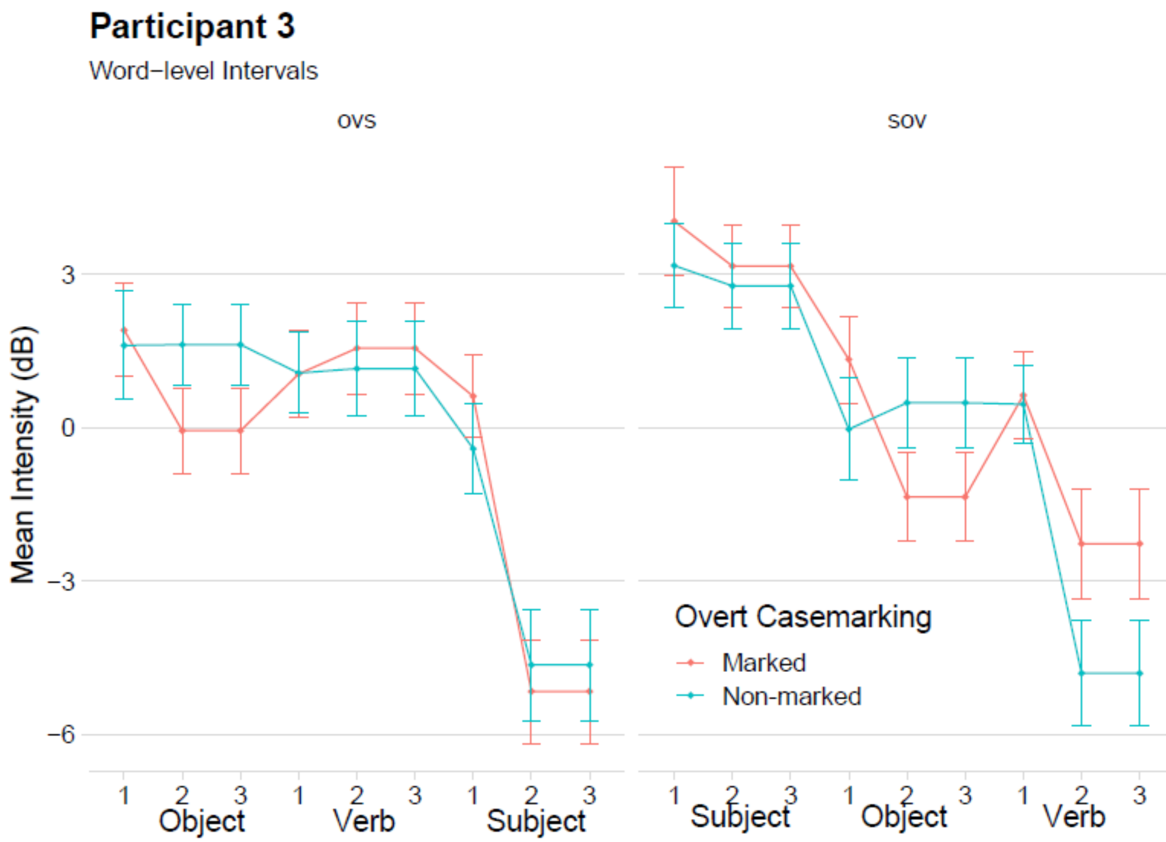

Participant 8
Word−level Intervals
ovs
sov
Mean F0 (Hz)
20
0
−20
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object Verb Subject
1 2 3 1 2 3 1 2 3
Subject Object Verb

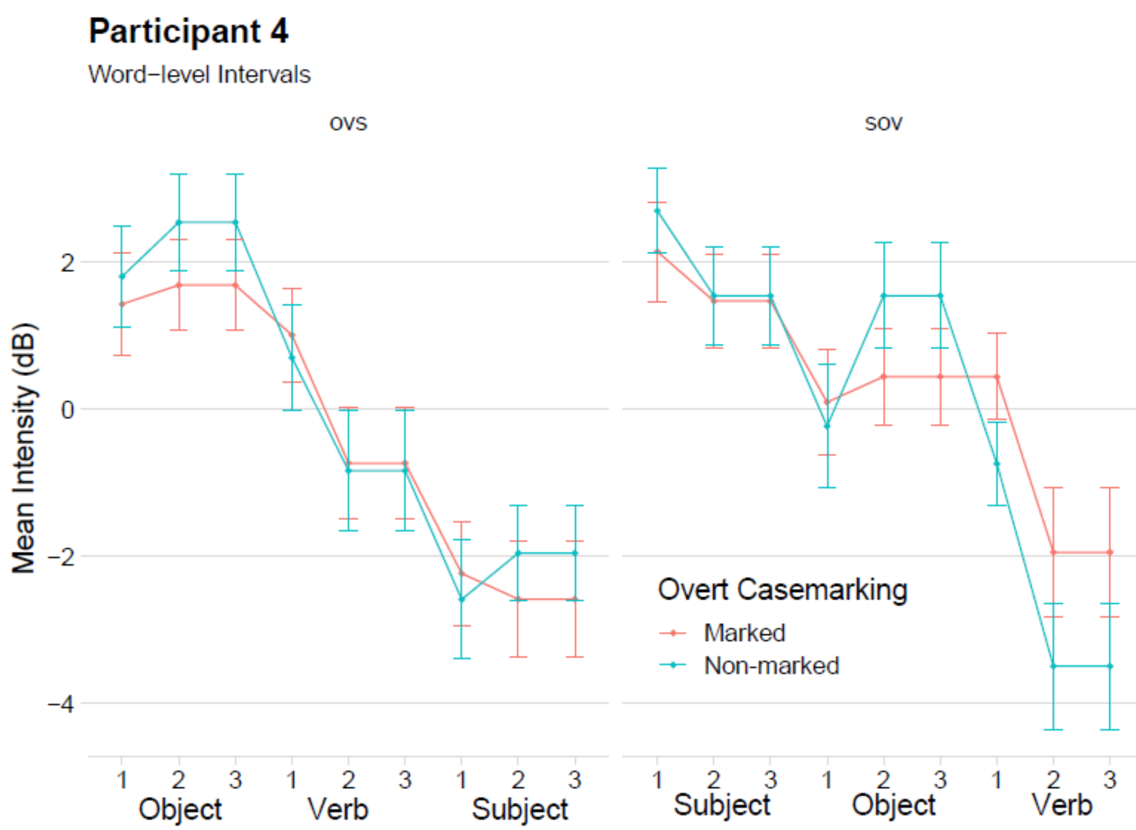

Participant 9
Word−level Intervals
ovs
sov
Mean F0 (Hz)
20
10
0
−10
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object Verb Subject
1 2 3 1 2 3 1 2 3
Subject Object Verb

Appendix D: Individual Participants' Intensity values for OVS sentences and SOV sentences.



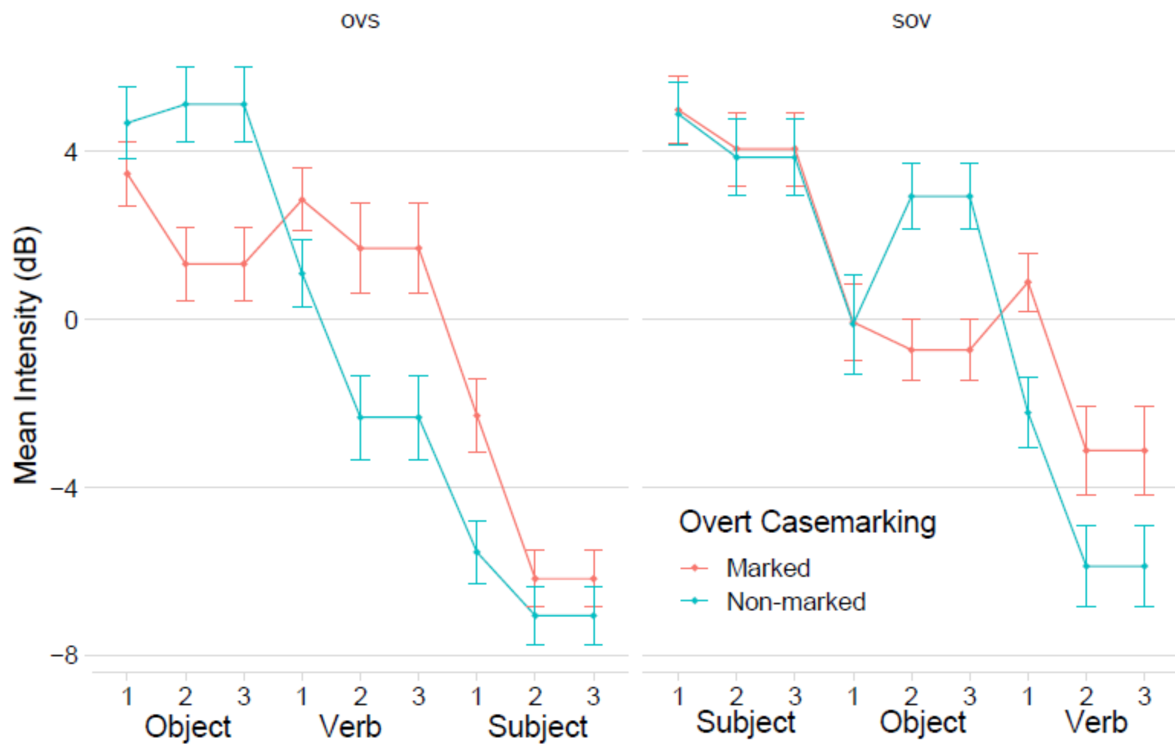



Participant 3
Word−level Intervals
ovs
sov
Mean Intensity (dB)
3
0
−3
−6
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object Verb Subject
1 2 3 1 2 3 1 2 3
Subject Object Verb


Participant 4
Word−level Intervals
ovs
sov
Mean Intensity (dB)
2
0
−2
−4
Overt Casemarking
Marked
Non-marked
1 2 3 1 2 3 1 2 3
Object Verb Subject
1 2 3 1 2 3 1 2 3
Subject Object Verb

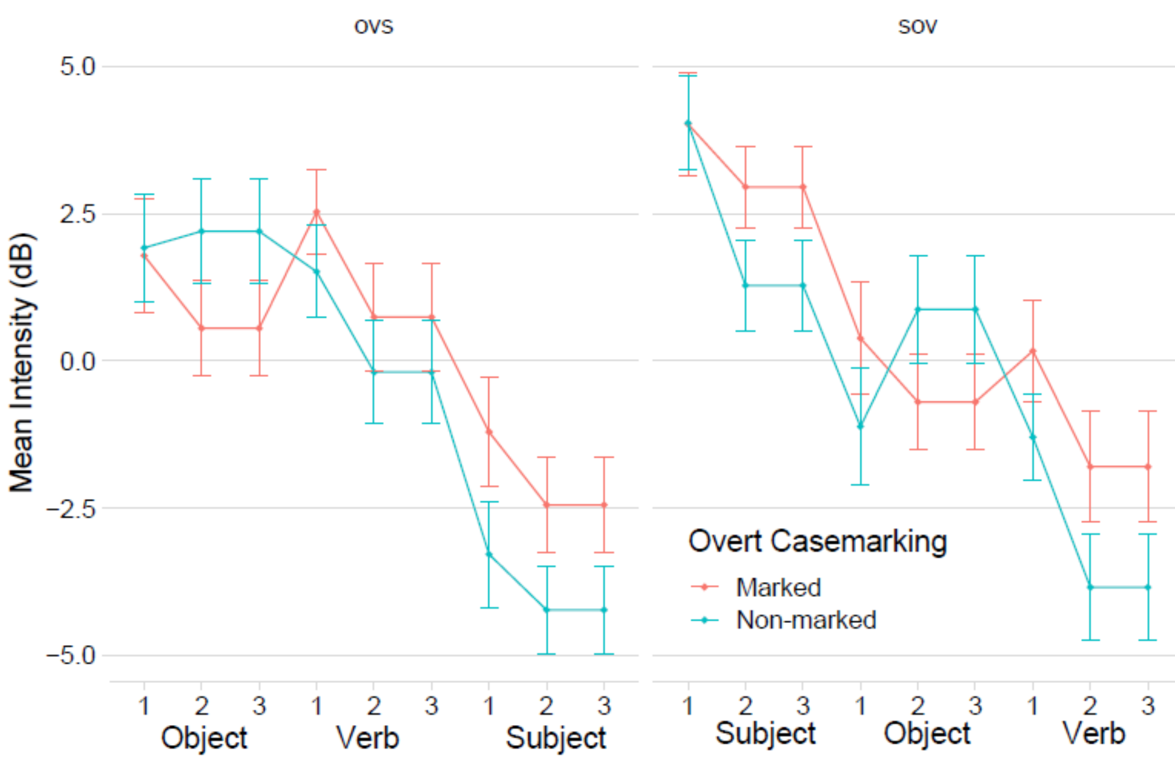
Participant 5

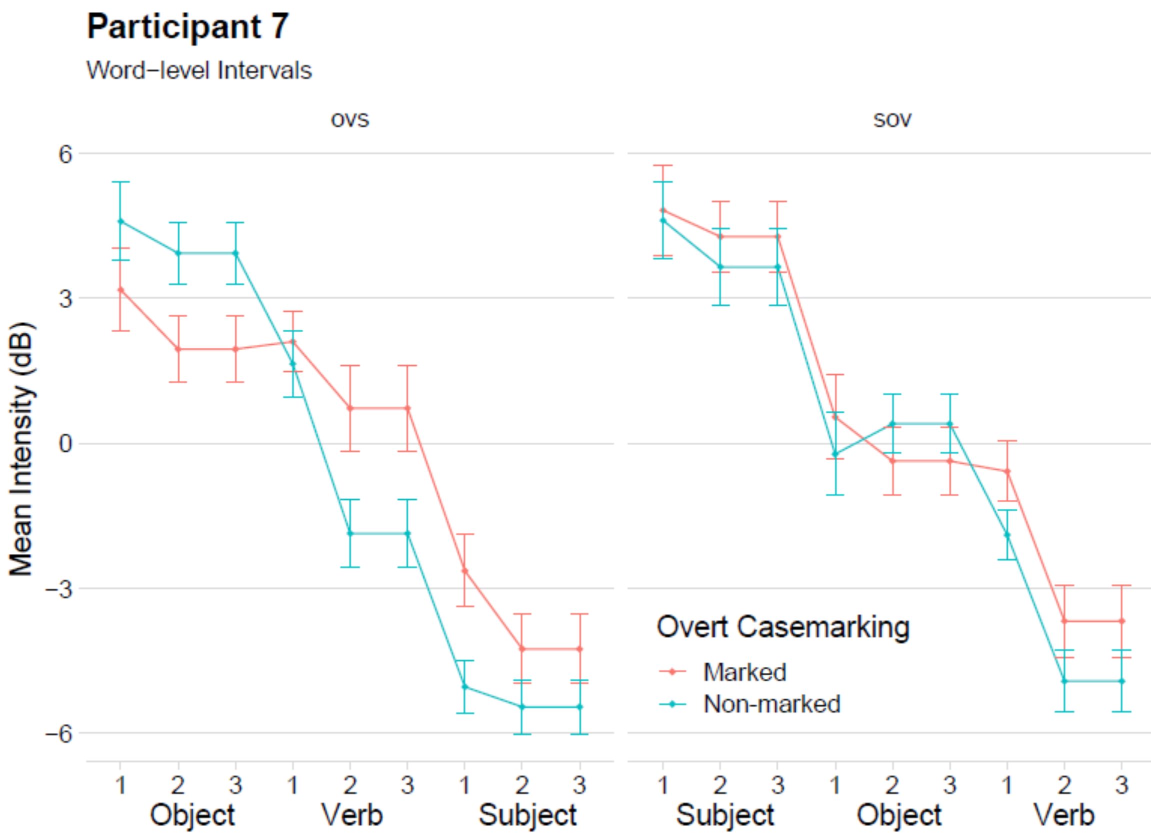
Word−level Intervals



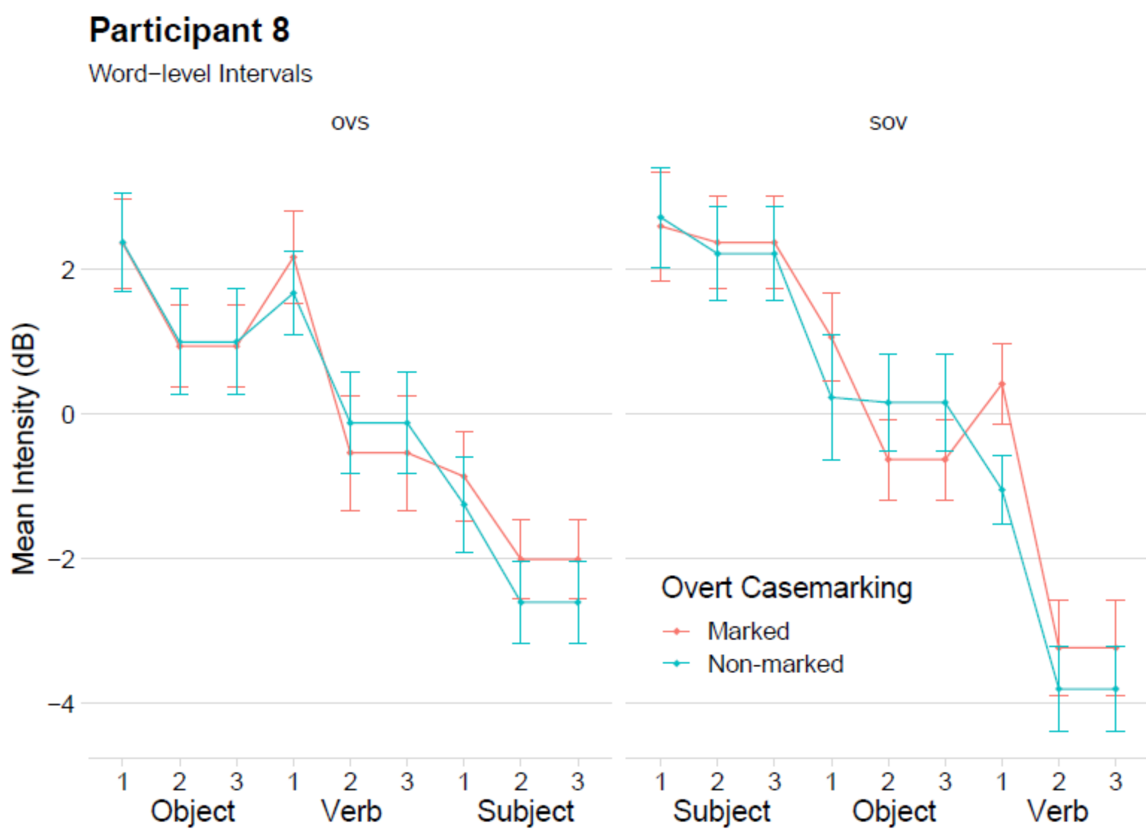
Participant 6

Word−level Intervals

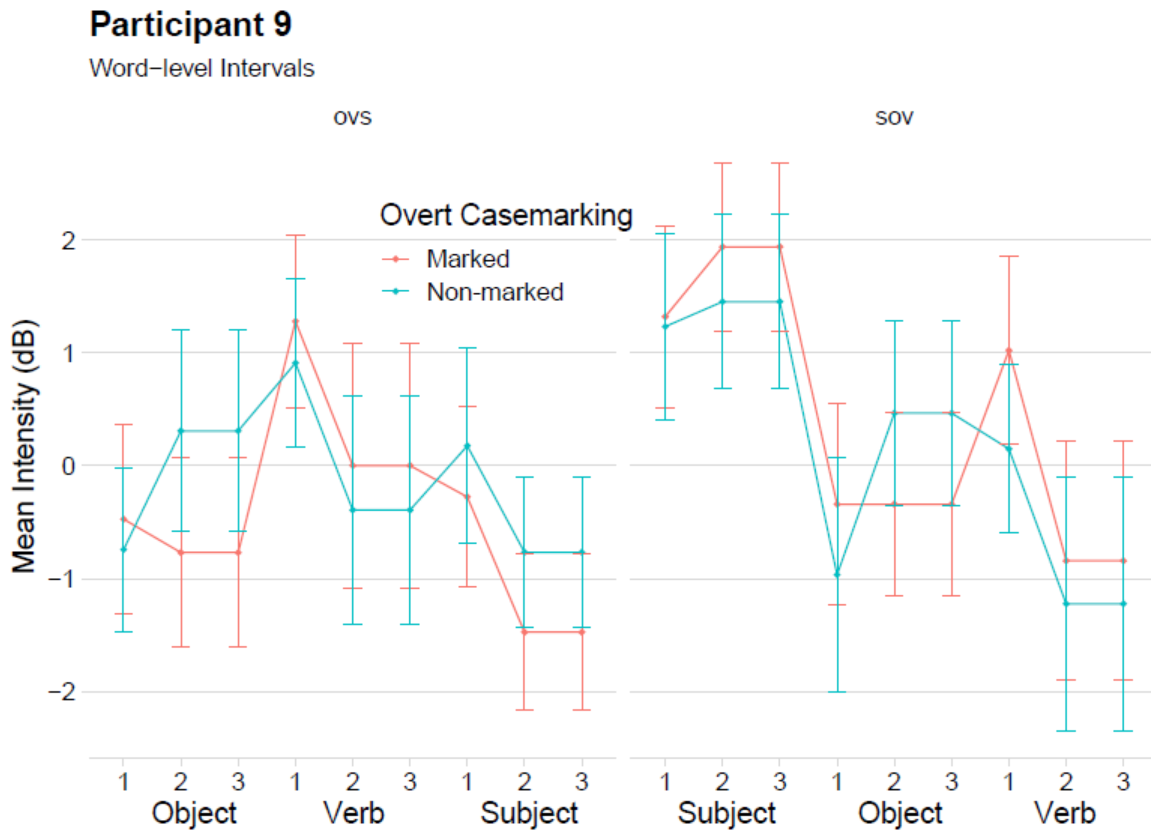




Participant 7
Word−level Intervals
ovs
sov
Mean Intensity (dB)
6
3
0
−3
−6
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
Subject
Object
Verb
Overt Casemarking
Marked
Non-marked



Participant 8
Word−level Intervals
ovs
sov
Mean Intensity (dB)
2
0
−2
−4
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
Subject
Object
Verb
Overt Casemarking
Marked
Non-marked


Participant 9
Word−level Intervals
ovs
sov
Overt Casemarking
Marked
Non-marked
Mean Intensity (dB)
2
1
0
−1
−2
1 2 3 1 2 3 1 2 3
Object
Verb
Subject
1 2 3 1 2 3 1 2 3
Subject
Object
Verb

Appendix E**:** Model specifications and tables for the four models used in analyses.

Random intercepts for participants and random slopes for the effects of grammatical role, the linear term, and the quadratic term were included in the models described in sections 3.1.2 and 3.2.2.
Random intercepts for scenarios and random slopes for the effects of case-marking, grammatical role and the linear term were included in the models.
Covariance between the effects of fixed effects in the error terms were fixed to be zero to reduce variance around the estimates.

Model: Normalized mean F0 by Case-marking (Sum coded with marked as 1 and non-marked as -1) * Grammatical Role (Sum coded with object as 1 and verb as -1) * (linear + quadratic contrasts for intervals, poly coded) * word order (Sum coded with OVS as 1 and SOV as -1) + (0 + Grammatical Role + linear + quadratic || participant) + (0 + case-marking + Grammatical Role + linear || scenario)

F0 Model (3.1.2)	Estimate ($\hat{\beta}$)	Standard Error	*t*-value	p-value
Intercept	-0.44	0.34	-1.27	.20
Interval (Linear, L)	-6.47	1.67	-3.85	<.005
Interval (Quadratic, Q)	4.25	1.02	4.14	<.005
Case-marking (C)	2.21	0.34	6.39	<.001
Grammatical Role (GR)	10.22	1.33	7.68	<.001
Word Order (WO)	5.43	0.34	15.68	<.001
L*C	-0.22	0.44	-0.50	.61
L*GR	8.97	0.44	20.55	<.001
L*WO	3.42	0.44	7.84	<.001
Q*C	0.32	0.44	0.74	.73
Q*GR	-1.35	0.44	-3.09	<.005
Q*WO	-0.78	0.44	-1.78	.07
C*GR	-4.18	0.30	-13.61	<.001
C*WO	-0.43	0.34	-1.27	.20
GR*WO	-3.90	0.30	-12.72	<.001
L*C*GR	-1.32	0.44	-3.02	<.005
L*C*WO	0.98	0.43	2.24	<.05
L*GR*WO	-0.30	0.43	-0.69	.48
Q*C*GR	0.93	0.44	2.13	<.05
Q*C*WO	-0.21	0.43	-0.48	.62
Q*GR*WO	-1.05	0.43	-2.42	<.05
C*GR*WO	.50	0.30	1.64	.10
L*C*GR*WO	.16	0.43	0.36	.71
Q*C*GR*WO	-0.46	0.43	-1.07	.28

Model: Normalized mean intensity by Case-marking (Sum coded with marked as 1 and non-marked as -1) * Grammatical Role (Sum coded with object as 1 and verb as -1) * (linear + quadratic contrasts for intervals, poly coded) * word order (Sum coded with OVS as 1 and SOV as -1) + (0 + Grammatical Role + linear + quadratic || participant) + (0 + case-marking + Grammatical Role + linear || scenario)

Intensity Model (3.2.2)	Estimate ($\hat{\beta}$)	Standard Error	*t*-value	p-value
Intercept	0.11	0.18	0.63	.52
Interval (Linear, L)	-0.69	0.19	-3.48	<.005
Interval (Quadratic, Q)	0.46	1.02	4.14	<.005
Case-marking (C)	0.14	0.18	0.79	.42
Grammatical Role (GR)	1.09	0.25	4.34	<.001
Word Order (WO)	1.02	0.18	5.50	<.001
L*C	-0.17	0.17	-1.03	.30
L*GR	0.70	0.06	10.32	<.001
L*WO	-0.04	0.17	-0.27	.78
Q*C	0.10	0.06	1.57	.11
Q*GR	-0.40	0.06	-5.96	<.001
Q*WO	-0.02	0.06	-0.31	.75
C*GR	-0.58	0.15	-3.73	<.001
C*WO	-0.10	0.18	-0.56	.56
GR*WO	-0.11	0.15	-0.73	.46
L*C*GR	-0.36	0.06	-5.33	<.001
L*C*WO	0.09	0.17	0.52	.59
L*GR*WO	-0.40	0.06	-5.99	<.001
Q*C*GR	0.21	0.06	3.07	<.005
Q*C*WO	-0.05	0.06	-0.85	.39
Q*GR*WO	0.23	0.06	3.46	<.001
C*GR*WO	0.08	0.15	0.52	.59
L*C*GR*WO	0.09	0.06	1.45	.14
Q*C*GR*WO	-0.05	0.06	-0.83	.40

For models described in 3.1.4 and 3.2.4:
Random intercepts for participants and random slopes for the effects of word order, the linear term, and the quadratic term were included in the model.
Random intercepts for scenarios and random slopes for the effects of case-marking and word order were included in the models.
Covariance between the effects of fixed effects in the error terms were fixed to be zero to reduce variance around the estimates.

Model: Normalized mean F0 by Case-marking (Sum coded with marked as 1 and non-marked as -1) * Word Order (Sum coded with OVS as 1 and SOV as -1) * (linear + quadratic contrasts for intervals, poly coded) + (0 + word order + linear + quadratic || participant) + (0 + case-marking + word order || scenario)

F0 Model (3.1.4)	Estimate ($\hat{\beta}$)	Standard Error	*t*-value	p-value
Intercept	0.08	0.47	0.18	.85
Interval (Linear, L)	4.42	4.55	0.97	.35
Interval (Quadratic, Q)	0.45	1.57	0.28	.78
Case-marking (C)	0.19	0.47	0.41	.68
Word Order (WO)	2.24	0.93	2.40	<.05
L*C	1.29	0.63	2.05	<.05
Q*C	0.95	0.63	1.51	.13
L*WO	1.34	0.63	2.12	<.05
Q*WO	0.62	0.63	.99	.32
C*WO	-2.14	0.47	-4.51	<.001
L*C*WO	-1.72	0.63	-2.72	<.01
Q*C*WO	-0.43	0.63	-0.68	.49

Model: Normalized mean intensity by Case-marking (Sum coded with marked as 1 and non-marked as -1) * Word Order (Sum coded with OVS as 1 and SOV as -1) * (linear + quadratic contrasts for intervals, poly coded) + (0 + word order + linear + quadratic || participant) + (0 + case-marking + word order || scenario)

Intensity Model (3.2.4)	Estimate ($\hat{\beta}$)	Standard Error	*t*-value	p-value
Intercept	0.16	0.21	0.76	.44
Interval (Linear, L)	-0.54	0.12	-4.45	<.005
Interval (Quadratic, Q)	0.31	0.10	3.08	<.005
Case-marking (C)	-0.19	0.21	-0.91	.35
Word Order (WO)	-0.37	0.24	-1.55	.12
L*C	-0.10	0.10	-1.04	.29
Q*C	0.06	0.10	0.60	.54
L*WO	.07	0.10	0.71	.47
Q*WO	-0.04	0.10	-0.41	.67
C*WO	-0.26	0.21	-1.25	.21
L*C*WO	-0.24	0.10	-2.36	<.05
Q*C*WO	0.14	0.10	1.36	<.001