Syntactic pausing? Re-examining the associations

> The right word may be effective, but no word was ever as effective as a rightly timed pause.
>
> – Mark Twain

1 Introduction

Linguists primarily conceptualise language as a continuous flow of speech, broken up by pauses. These pauses can be thought of as temporary 'breaks' in which the transmission of content between interlocutors is cessated. They are typically understood as either filled or silent, depending on whether interlocutors perceive the break to contain a voiced section (e.g. 'um', 'uh' in English) or silence (Belz & Trouvain 2019, Clark & Fox Tree 2002, Zellner 1994, see also Goh et al. 2023). This paper focuses on silent pauses, which go beyond the pure physiological need to breathe; rather, pausing "operates in happy synchrony with some basic functional segmentations of discourse" (Chafe 1994: 57).

The use of silent pauses has been studied from various perspectives in linguistics. From a psycholinguistic angle, silent pauses have been examined for their function of marking prosodic structures which support syntactic parsing and disambiguation, especially in early L1 acquisition (e.g. Christophe et al. 2008, de Carvalho et al. 2017, Hawthorne & Gerken 2014, Speer & Ito 2009). Much theoretical work on the prosody-syntax interface has similarly focused on the association of silent pauses with syntactic structure, notably the co-occurrence of silent pauses with syntactic boundaries (e.g. Goldman-Eisler 1972, Nespor & Vogel 2007, Selkirk 1984, 2011, Truckenbrodt 1995, 2007, Watson & Gibson 2004). Conversation analysis and related disciplines, on the other hand, have examined the use of silent pauses as markers of turn-taking in conversational discourse (e.g. Sacks et al. 1974, Taboada 2006, Weilhammer & Rabold 2003, Zellner 1994). The role of pausing in the perception of fluency has also been investigated (e.g. Belz et al. 2017, Bosker et al. 2013, de Jong 2016, Kahng 2018), typically comparing differences in pausing between native and non-native speakers. Silent pauses have also been studied for their contribution to rhetorical style, as well as to sociolinguistic and

cross-cultural variation (e.g. Duez 1982, Kendall 2013, Schleef 2019, Šturm & Volín 2023, Tannen 1985, Walker 1985).

In this paper, we concentrate on the association between between syntactic boundaries and pause location as well as duration. Using spontaneous speech data from seven typologically distinct languages, we explore the distribution of silent pauses to see whether their location and length correlate with higher-level syntactic groupings, namely main and dependent clause boundaries. Our results point to a tendency for main clause boundaries to be accompanied by silent pauses across languages, while dependent clause boundaries are less likely to co-occur with pauses. Our results confirm previous findings in that pauses at clause boundaries tend to be longer than pauses within clauses across all languages in our dataset.

The paper is structured as follows. In Section 2 we provide an overview of previous research into the relation of pausing and clause boundaries. We then introduce our data, annotation and methodology in Section 3. Section 4 presents the results of the present study; we discuss and contextualise them within the wider debate around pausing in Section 5. Section 6 concludes.

2 Silent pauses and syntactic structure

2.1 Pauses and the syntax-prosody interface

Silent pauses as markers of boundaries have long been observed to pattern with syntactic boundaries, especially with clause boundaries (e.g. Cooper & Paccia-Cooper 1980, Goldman-Eisler 1968, Nespor & Vogel 2007). However, pauses do not necessarily co-occur with syntactic boundaries, i.e. they can be observed outside of the syntactically determined positions. Such pauses are usually attributed to performance, e.g. to hesitation, processing, style (cf. Nespor & Vogel 2007: 189, 219) as well as breathing, especially in faster speech (cf. Grosjean & Collins 1979). In addition, we find certain types of syntax-prosody mismatches besides performance-related pauses (Shattuck-Hufnagel & Turk 1996). Syntactic structures are not always sufficient to predict the prosodic structure of an utterance, and several prosodic realisations of the same syntactic unit are possible.

There are thus various reasons to assume a certain degree of independence between syntactic and prosodic structures, while allowing for a high degree of interaction at the same time. Such a position on the relation between syntax and prosody is captured by approaches which assume an intermediary prosodic structure to which morphosyntax is mapped (cf. Nespor & Vogel 2007, Selkirk 1984, 2011). Our study follows this line of research, using crosslinguistic corpus data of spontaneous speech to examine to what extent pauses are

associated with syntactic boundary positions.

2.2 Pause types and functions

There are different ways that silent pauses can be distinguished, for instance according to where they occur. Previous studies have distinguished between pauses at sentence, clause, and phrase boundaries as well as pauses that do not coincide with any syntactic boundaries. Example 1 illustrates different possible locations of silent pauses in spontaneous spoken English, taken from DoReCo (Seifart et al. 2022). Here, each line corresponds to a prosodic unit delimited by a silent pause whose length is indicated in brackets.

(1) Silent pauses in spontaneous spoken English
(0298-0300_DoReCo_doreco_sout3282_mc_english_kent02_b)

a. so I messed the skin up (1430ms)
b. it wadn't no good then (670ms)
c. cause I was close to him you see blowed [false start] (70ms)
d. a great hole in him (2503ms)
e. catched a deer in a (645ms)
f. snare one day I went down there as I told you about how I always […]

The pauses following (1a) and (1d) coincide with main clause boundaries . The pause at the end of (1b) occurs at a dependent clause boundary. The three remaining rows include pauses that do not coincide with clausal or phrasal boundaries. In (1c), a pause occurs between a transitive verb and its object, where it follows a false start. A similar situation can be seen in the transition between (1e) and (1f), where a pause separates a determiner from its head noun.

Silent pauses such as those in (1a), (1b) and (1d) are typically understood within psycholinguistics as helping to facilitate the disambiguation of syntactic units. This has been shown for both adult L1 language use (e.g. Frazier et al. 2006, Kjelgaard & Speer 1999, Petrone et al. 2017, Schafer et al. 2000) as well as L1 and L2 language acquisition (e.g. Christophe et al. 2008, Goad et al. 2021, Hawthorne & Gerken 2014, Speer & Ito 2009). At the same time, these pauses also help the speaker to process the previous utterance as well as prepare and plan further utterances (e.g. Cooper & Paccia-Cooper 1980, Ferreira 1991, Fuchs et al. 2013, Goldman-Eisler 1968, Krivokapić 2007, Krivokapić et al. 2020). They can also be used for interactional purposes, such as holding the floor during a conversation (Levelt 1993, Maclay & Osgood 1959, Wennerstrom & Siegel 2003).

Other pauses, such as the clause-internal ones in (1c), (1e) and (1f), are argued to result from high processing demands in situations in which speech planning is comparatively difficult. Planning difficulties can be caused by various factors, e.g. long or complex following syntactic or prosodic structures, or difficulties in lexical access with low-probability items (cf. Beattie & Butterworth 1979, Belz et al. 2017, Goldman-Eisler 1961a, Hartsuiker & Notebaert 2010).

As we work with corpus data in the present study, we cannot probe for cognitive reasons behind the distribution of pauses. As such, we do not assume that pauses occur for hesitation, planning, or stylistic reasons. We instead distinguish between three types of pauses based on their location in relation to syntactic boundaries: pauses within clauses, pauses at main clause boundaries and pauses at dependent clause boundaries (see Section 3 for more details and examples). This distinction allows us to measure the association of different clausal contexts with pauses, and to assess how robust the patterns found are across languages (cf. Section 4.2).

2.3 Pause duration

Studies into the distribution of pause duration often show a distinction between brief pauses and longer pauses, with longer pauses typically found at sentence boundaries and shorter pauses at dependent clause boundaries and within clauses. In an early study, Goldman-Eisler (1972) reports for spontaneous spoken English data that pauses between sentences mostly have durations above 500ms, while pauses between clauses (within sentences) tend to be shorter than 500ms. Grosjean & Deschamps (1972) report comparable pause durations for other French and English datasets; Fletcher et al. (2004) similarly group pauses in Dalabon (Gunwinyguan, Australia) into medium (500ms) and long (>1000ms) pauses.

Pause distribution and duration in spontaneous French datasets show similar trends, even when more detailed distinctions of pause type are made. Duez (1982: 24) finds that pauses within phrases (401ms) and within clauses (632ms) tend to be shorter than pauses between clauses (802ms) in casual interviews. In a more recent study, Candea (2000) reports that pauses at sentence boundaries have average durations of 900ms, while pauses at clause boundaries within sentences average at 600ms (Candea 2000: 168). In contrast, hesitation pauses within clauses are reported to have an average duration of 560ms (Candea 2000: 181).[1] Campione & Véronis (2002) focus on pause duration in French spontaneous speech data and read speech from five European languages. For the spontaneous data, they propose a more fine-grained distinction than previous studies, distinguishing between brief (<200ms),

[1]Candea (2000) only classifies silent pauses as hesitation pauses if they co-occur with another hesitation marker.

medium (200-1000ms), and long (>1000ms) pauses.

While the exact details differ across studies, a robust picture emerges with pauses at clause boundaries being systematically longer than pauses within clauses (also see Yang 2004).[2] We assess to what extent this association of pause location and duration is reflected in our data, how it relates to previous findings, and how robust the patterns are across languages (cf. Section 4.3).

3 Data and annotation

3.1 Dataset

In our study, we investigated how pauses were distributed across seven typologically distinct languages in the Multi-CAST corpora (Haig & Schnell 2021).

language	N speaker	N utterances	genre
Arta (Austronesian) (Kimoto 2019)	3	227	AN, TN
Nafsan (Austronesian) (Thieberger & Brickell 2019)	3	163	TN
Teop (Austronesian) (Mosel & Schnell 2015)	4	1019	TN
Tondano (Austronesian) (Brickell 2016)	6	1254	SN, AN
Mandarin (Sino-Tibetan) (Vollmer 2020)	3	845	TN
Tabasaran (Nakh-Daghestanian) (Bogomolova et al. 2021)	2	629	TN, AN
Northern Kurdish (Indo-European) (Haig et al. 2015)	2	555	TN

Table 1: Overview of the dataset.

[2]Other factors to influence pause duration are the complexity of preceding/following syntactic and prosodic units (Ferreira 1993, Gee & Grosjean 1983, Goldman-Eisler 1961b, Krivokapić 2007), speech rate, speaking style, genre as well as semantic factors (Grosjean et al. 1979, Kendall 2013, Klatt 1976, Yang 2004, Zellner 1994).

The languages are listed in Table 1 together with information on the number of different speakers in the corpus, the number of utterances, and the genres of texts used. We chose a subset of data which consisted of realistic monologues, i.e. autobiographical narratives (AN) and traditional narratives (TN). Only Tondano includes one additional stimulus-based narrative (SN). As can be seen in Table 1, each corpus ranges from 160 to 1300 utterances per language from 2 to 6 different speakers. We used the .wav files along with the accompanying annotations in .eaf format provided in Multi-CAST (Haig & Schnell 2021).

3.2 Clause annotation

The majority of spontaneous speech corpora are created for language documentation purposes, and as such, differ as to how elaborate their morphosyntactic annotation is. Information on clausal boundaries below the utterance level is usually not explicitly annotated for, Multi-CAST being an important exception.

We took the annotation of clausal boundaries from GRAID (Grammatical Relations and Animacy in Discourse) annotations in Multi-CAST. In GRAID, a clause is generally defined as a predicate with its arguments, with more language-specific criteria for how to deal with, e.g., multi-verb predicates (Haig & Schnell 2014: 45). Main clauses are signalled with an annotation of "**##**" at their left edge, while dependent clauses are annotated as "**#**" at the left edge. The end of a dependent clause is generally not marked explicitly, as it usually coincides with the beginning of a new main clause. However, the end of a center-embedded dependent clause that does not correspond to the end of a main clause is additionally signalled by "**%**" in GRAID.

Examples (2) to (4) show how clause boundaries are annotated in GRAID (see Haig & Schnell (2014: §2.6, §4.1) for more details).[3] Example (2) consists of a main clause in Mandarin, marked as such by the initial "**##**".[4]

(2) Mandarin (mandarin_lzh_0007-0008)

ránhòu yě jiù shì súchēng=de zhù yuánwài **##**
then also ADV COP popular_name=MOD Zhu landlord **##**
'He was called Zhu landlord by people.'

In Example (3), we see an utterance from the Northern Kurdish corpus, consisting of a main clause with a dependent relative clause. In this case, the end of the relative clause

[3]For the sake of simplicity, we combined the GRAID annotations of clausal boundaries with the morphosyntactic glosses. In the original Multi-CAST annotations, these are two different annotation levels. The Multi-CAST data also contains detailed referential annotations which we do not include here, as they are not relevant for the purposes of the present study.

[4]For the sake of completeness, we added the initial main clause boundary marker "**##**" from the following clause for all examples.

corresponds to end of the main clause and does not receive an explicit annotation.

(3) Northern Kurdish (n_kurd_muserz01_003-004)

kur-ek-î wî hebû-ye # nav-ê kur-ê wî Mihemed
son-INDF-EZ 3SG.OBL exist.PST-PRF.3SG # name-EZ son-EZ 3SG.OBL Mihemed

bû-ye ##
COP.PST-PRF.3SG ##
'He had a son, whose name was Mihemed.'

The Nafsan example in (4) shows a somewhat more complex utterance. It is made up of three main clauses, the first of which contains an additional dependent clause. Note that the end of the dependent clause does not coincide with the end of the main clause, which is why it is additionally marked with the annotation by "%" in (4a).

(4) Nafsan (nafsan_tafra_0001)

a. ## # selwan tu=paakor nametp̃ag ntau % ra=to tu teesa
while 1PL.IN.RS=arrive end year % 1DU.EX.RS=HABIT give child

tete nanromien
some present
'When we got to the end of the year we would give the children a present …'

b. ## ru=to ni apu go atien negar-wes
3PL.RS=stay with grandfather and grandmother 3PL.POSS-3SG.OBL
'… for them to take to their Apu and Ati …'

c. ## nanromien sees pan tu-e-r ki-n Ertap ##
present small go give-TS-3PL.O PREP-DST Eratap
'… a small present they could give to them at Eratap.'

We used this clause information from the original annotation to distinguish between main clause boundaries (`cb_main`), dependent clause boundaries (`cb_dep`) and clause-internal boundaries (`no_cb`). Main and dependent clause boundaries as shown in Examples (2) to (4) were directly taken from the annotation. The remaining boundaries were classified as clause-internal.

3.3 Pause annotation

Spontaneous speech corpora often do not include annotations for phonological or prosodic properties.[5] As such, the re-use of these corpora for this paper required us to additionally

[5]An important exception to this is the collection of "Documentation Reference Corpora (DoReCo)" (Seifart et al. 2022), which consists of spontaneous speech corpora of 51 typologically distinct languages with phonetic-level annotation.

annotate silent pauses in order to explore any potential associations between pauses and clause boundaries. The pause annotation process we followed is described in detail in the supplementary document "data-extraction-processing.pdf"; we provide a brief summary here and some illustrative examples for the reader's orientation.[6]

We firstly annotated silent pauses using a silence recogniser and followed this with a round of manual checks and corrections. This resulted in an annotation of pause and speech segments. We then classified both pause and speech segments into one of three types (`cb_main`, `cb_dep`, `no_cb`) according to whether or not a main or dependent clause boundary fell into their duration. This was done automatically comparing the time stamps of the pre-existing clause boundary annotations and those of the newly-annotated silent pause and speech segments. Since the Multi-CAST data is only aligned with the acoustic signal on the utterance level and not on the clausal level, we could not automatically specify the location of a pause occurring within a clause further.[7]

Tables 2 to 6 show examples of the resulting pause annotation for examples (2) to (4) from Section 3.2. The first row shows the utterance including the original clause boundary annotation from Multi-CAST. The second *pause* row then shows whether or not a silent pause was detected. If a pause was detected, the *type* row indicates which type of pause we are dealing with, and the *dur* row shows its duration. In the Mandarin example in Table 2, we see that both the initial and final clause boundaries coincide with a pause, which we annotated as `cb_main`. In addition, we find another silent pause occurring within the clause, which was marked as `no_cb`.

utt	##	ránhòu yě jiù shì súchēng=de zhù yuánwài	##
pause	✓	✓	✓
type	cb_main	no_cb	cb_main
dur	124ms	235ms	226ms

Table 2: Pause annotation for Mandarin example (2)

In the Northern Kurdish utterance in Table 3, the main clause boundaries also coincide with silent pauses. The dependent clause boundary between the two main clauses, however, is not accompanied by a silent pause. In addition, no pauses occur within the two clauses shown

[6]The final, processed dataset used for the analyses in this paper can be found in the supplementary materials ("data-pausing.csv").

[7]Annotating for the exact location of all `no_cb` pauses would have required a significant manual annotation effort. This would have gone beyond the purposes of the present paper, which focuses on the co-occurrence of silent pauses and clausal boundaries. However, see Section 4 in the Supplementary Materials for a comparison to DoReCo pause annotations.

in Table 3.

utt	##	kur-ek-î wî hebû-ye	#	nav-ê kur-ê wî Mihemed bû-ye	##
pause	✓	✗	✗	✗	✓
type	cb_main				cb_main
dur	2343ms				948ms

Table 3: Pause annotation for Northern Kurdish example (3)

The Nafsan example from (4) is shown in Tables 4 to 6. Table 4 contains the first main clause together with its dependent clause. Here, silent pauses only occur at the end of the dependent clause (cb_dep) and within the second part of the inherited main clause annotation (cb_main).

utt	## #	selwan tu=paakor nametp̃ag ntau	%	ra=to tu teesa tete nanromien
pause	✗	✗	✓	✓
type			cb_dep	no_cb
dur			1790ms	482ms

Table 4: Pause annotation for Nafsan example (4a)

utt	##	ru=to ni apu go atien negar-wes
pause	✓	✗
type	cb_main	
dur	1125ms	

Table 5: Pause annotation for Nafsan example (4b)

utt	##	nanromien sees pan tu-e-r ki-n Ertap	##
pause	✓	✓	✓
type	cb_main	no_cb	cb_main
dur	1692ms	533ms	1237ms

Table 6: Pause annotation for Nafsan example (4c)

Table 5 shows that another silent pause coincides with the boundary between the two main clauses in Table 4 and 5. Finally, we see in Table 6 that both clausal boundaries co-occur with a pause (`cb_main`), and that the main clause itself also contains a silent pause (`no_cb`).

4 Results

4.1 Overall Tendencies

We first inspect the distribution of clause boundaries. Table 7 shows how the three clausal contexts are distributed in the seven languages. Pause/speech segments with no clause boundaries (`no_cb`) make up around 55-70% of annotations in all languages, with an average of 61% across the dataset. The proportions of `cb_main` and especially `cb_dep` contexts are more language-dependent. For example, dependent clause boundaries make up 23% of all boundaries in Northern Kurdish, while they only amount to 4% in Mandarin.

Table 7: Proportion of clausal contexts.

Language	`no_cb`	`cb_main`	`cb_dep`	N segments
Arta	0.66	0.20	0.14	1205
Mandarin	0.62	0.33	0.04	3259
Nafsan	0.61	0.27	0.11	1716
Northern Kurdish	0.56	0.21	0.23	4112
Tabarasan	0.57	0.22	0.21	3798
Teop	0.59	0.33	0.08	3156
Tondano	0.70	0.24	0.06	3717
Average	0.61	0.26	0.13	

Figure 1 builds on the distribution of clausal contexts shown in Table 7 by including information about how they relate to pause and speech segments in the seven languages. Comparing the distribution of the three clausal contexts in pause and speech segments (left vs. right bars), we see clearly that pause segments contain a higher proportion of clause boundaries than speech segments in all languages. The exact proportions differ across languages, with clausal boundaries co-occurring with 75% of all pauses in Northern Kurdish but with only around 40% of all pauses in Tondano.

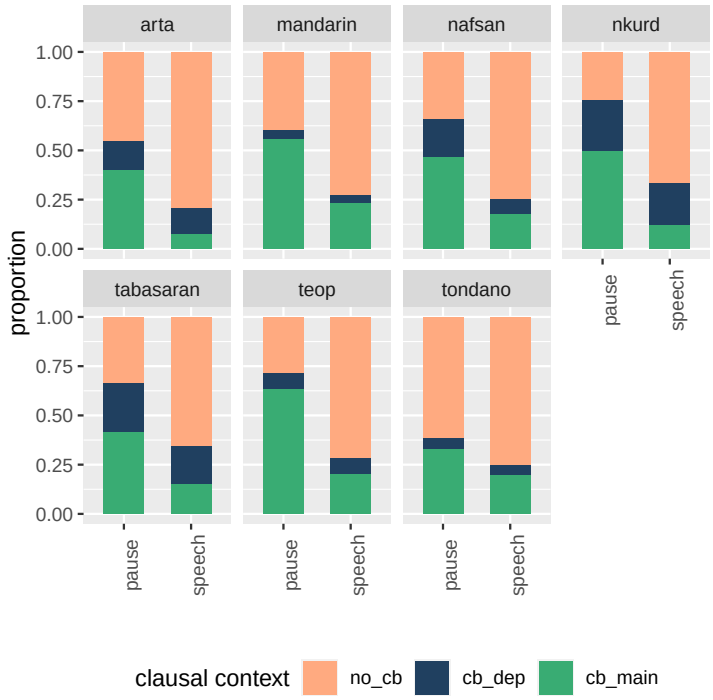

Figure 1: The raw distribution of clausal contexts across speech and pause segments.

The overall distribution of clausal contexts across our dataset suggest that clause boundaries more often co-occur with a pause than with speech. However, in Tondano and Arta, pauses are more likely to occur within clauses than at a clause boundary. The probability of pauses co-occurring with dependent clause boundaries is also substantially lower in Tondano, Mandarin, and Teop, likely due to the overall low proportion of dependent clause boundaries in these four languages to begin with (cf. Table 7). This suggests that the presence of a clausal boundary does not necessarily entail the occurrence of a silent pause and that the association between clause boundaries and pauses is indeed more complex than a simple one-to-one relationship.

4.2 The association of clause boundaries with pauses

The first question that this paper addresses is to what extent the different clausal contexts are associated with silent pauses. To do this, we fitted a Bayesian logistic regression model to assess the probability of pauses across the three types of clausal contexts. We added varying intercepts as well as varying slopes over clausal contexts for individual speakers.[8]

[8]The model formula is: pause ~ 1 + cl_context * language + (1 + cl_context|speaker). The code for the model is documented in "code-analysis.r" in the supplementary materials.

The model was fitted using Bayesian methods with Stan (Carpenter et al. 2017) and brms (Bürkner 2017) in R (R Core Team 2021).

Figure 2 shows the conditional effects of clausal contexts on the probabilities of pauses. The points in Figure 2 correspond to the mean of the probability density of the predictions; the whiskers correspond to the 95% uncertainty intervals. This means that we can be 95% confident that the actual probability will lie within this interval based on the data and the model.

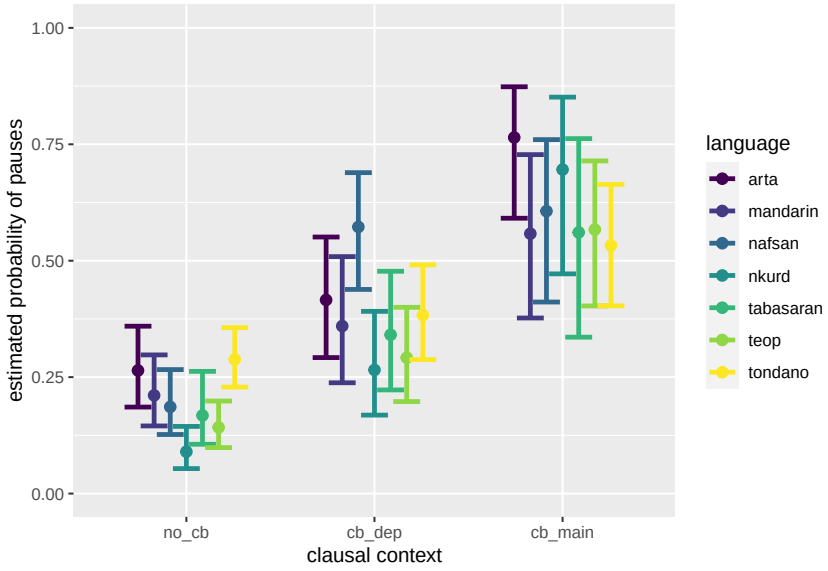

Figure 2: Predicted probability of pauses across clausal contexts.

While we see a certain degree of variation across the seven languages, the overall pattern is robust. The probability of pauses occurring gradually increases from environments within clauses (`no_cb`) to dependent clause boundaries (`cb_dep`) and main clause boundaries (`cb_main`). For main clause boundaries, the probability of co-occurrence with pauses is between 0.5 and 0.75 in all languages, the lowest prediction being 0.53 for Tondano and the highest being 0.76 for Arta. For dependent clause boundaries, the predicted co-occurrence probabilities for all languages are slightly lower, ranging from 0.29 in Teop to 0.57 in Nafsan. For the no clause boundary context, we find the overall lowest probability of pause co-occurrence, from 0.14 in Teop to 0.28 in Tondano. We furthermore generally find a higher level of cross-linguistic variation in the `cb_main` and `cb_dep` contexts than in the `no_cb` condition.

The robust patterning of pause co-occurrence probabilities across our dataset illustrates a clear trend. As the level of the syntactic boundary juncture increases, so too does the overall predicted probability of a pause co-occurring. Looking at the uncertainty intervals, we find

that they overlap between "neighbouring" categories, i.e. within clauses and dependent clause boundaries, as well as dependent and main clause boundaries. Only Nafsan differs in that both types of clause boundaries have a similarly higher probability of pauses compared to boundaries within clauses. Still, the overall results suggest that there is no crosslinguistically robust, categorical distinction either between main clause boundaries and other contexts or between non-clausal boundaries and clausal boundaries. There is, however, a clear difference between `no_cb` and `cb_main` in terms of pause probability in all seven languages.

4.3 Pause durations

The second question that this paper aims to address concerns pause durations in relation to their location. Figure 3 shows the raw distribution of pause durations across the seven languages of the dataset. The general trends are fairly robust. The majority of silent pauses in the dataset have a duration up to 750ms, few pauses are longer than 1000ms, and pauses longer than 2500ms are rare. However, Arta, Nafsan and Northern Kurdish have a somewhat larger portion of silent pauses longer than 750ms.

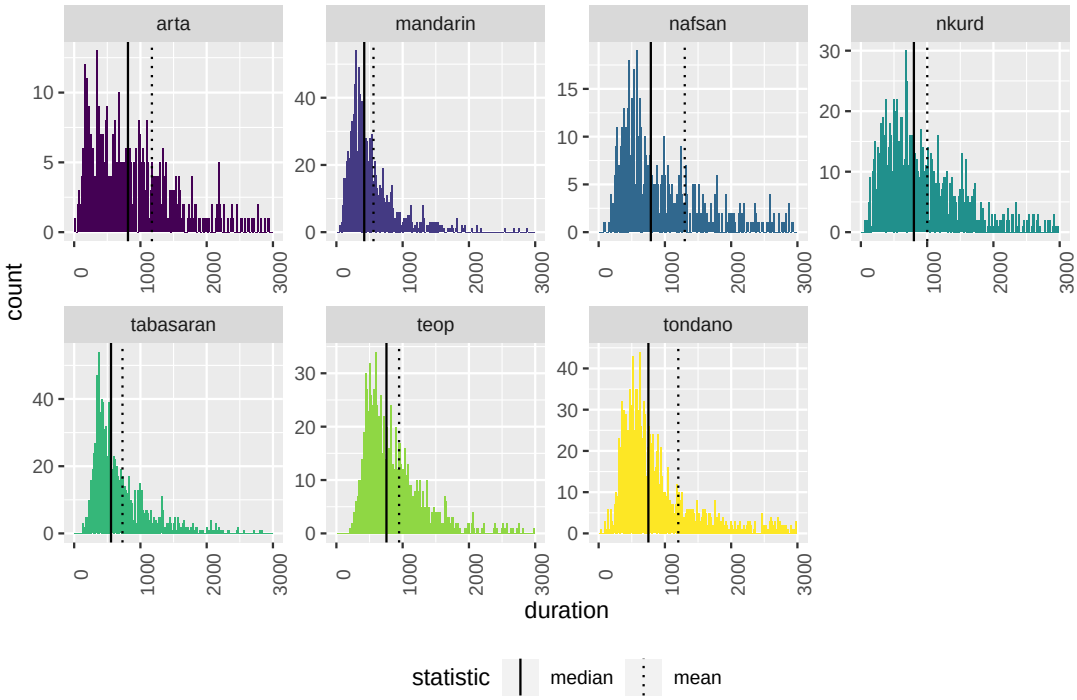

Figure 3: Observed pause duration by language.

Furthermore, the average pause duration is more or less comparable across languages. As can be seen in Figure 3, the median pause duration (the solid vertical line) is between 500ms

and 1000ms for all languages, apart from Mandarin.However, we see more variation of the mean pause duration (dotted black line) across languages. Arta, Nafsan, and Tondano show a larger difference between median and mean pause duration, as these three languages have more longer pauses than the other languages of the dataset. Stil, we can say that the distribution of pause length, albeit not identical, is comparable across the seven languages.

Figure 4 shows the distribution of pause duration together with median and mean values across the three clausal contexts. These distributions along with the median and mean suggest a difference in pause duration associated with the absence or presence of a clause boundary. Most pauses within clauses have a length of below 750ms, with only a few pauses longer than 750ms and even fewer above 1000ms. In contrast, we find quite a substantial proportion of pauses with durations above 1000ms up to 3000ms for both types of clause boundaries. This binary difference between pauses within clauses and pauses at clausal boundaries is further reflected in their median pause durations: 537ms for `no_cb` vs. 796ms and 788ms for `cb_dep` and `cb_main`, respectively.

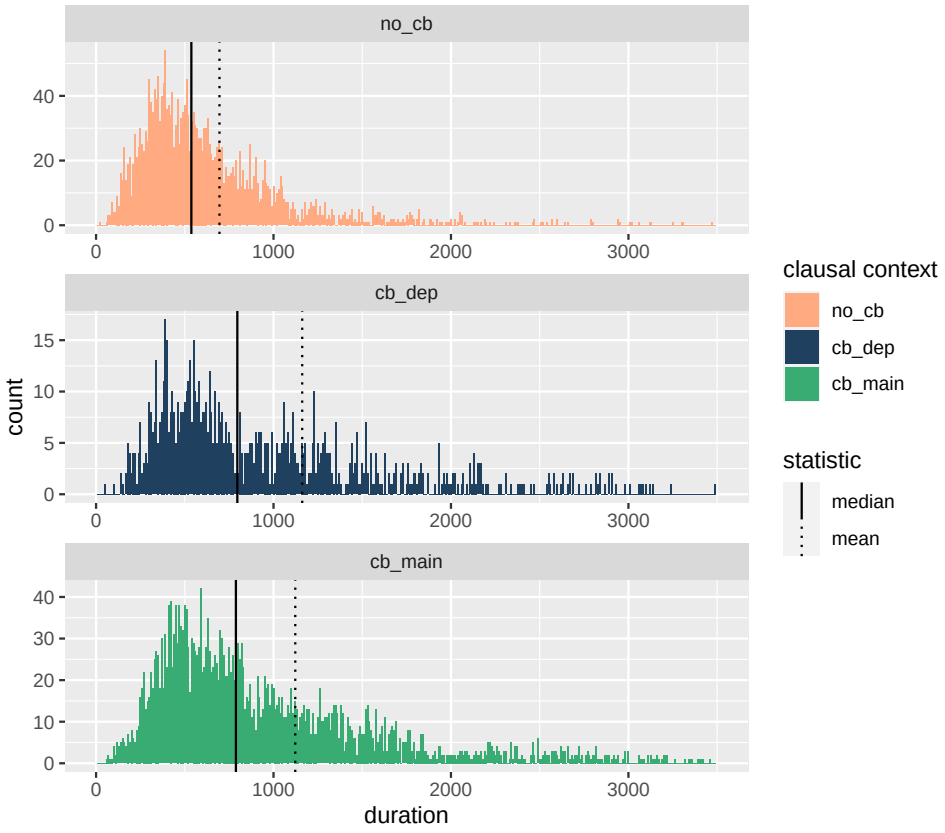

Figure 4: Observed pause duration by clausal context.

To test the robustness of the observations from the raw distributions, we fitted a Bayesian

regression model to predict the duration of pauses from clausal contexts and languages while controlling for the effects of single speakers.[9] Figure 5 shows the conditional effects of clausal contexts and languages on pause durations.

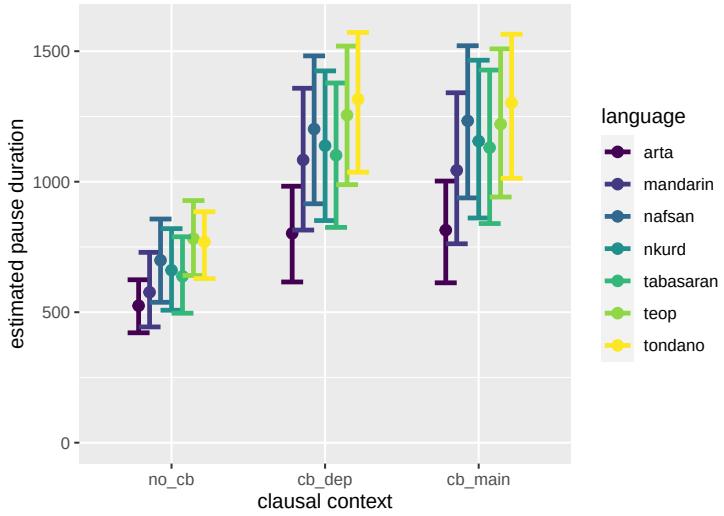

Figure 5: Estimated pause duration by clausal context.

The model predictions support the hypothesis of a two-way distinction of pauses based on their duration. Pause durations are predicted to be between 524ms and 781ms very consistently across languages when no clause boundary is present; the comparatively small credible intervals show that there is less variation of pause durations uncounted for. At main and dependent clause boundaries, the average predicted pause duration varies to a greater extent within and across languages. Pauses are predicted to be between 1000ms and 1300ms for both types of clausal boundaries in all languages except for Arta. Arta exhibits shorter pause durations in all three contexts compared to the other languages. However, the relative difference between the boundaries within clauses (525ms) and clausal boundaries (802ms and 814ms) is maintained. We should be careful before concluding that this indeed reflects a real crosslinguistic difference, as the Arta subcorpus is comparatively small (cf. Table 1). Larger datasets with data from more speakers are needed to be more certain about what seems to be a language-specific preference towards shorter pauses in our data.

[9]The model formula is: duration ~ cl_context * language + (1 + cl_context|speaker). The code for the model is documented in "code-analysis.r" in the supplementary materials.

5 Discussion

Our results evidence a complex interaction between pausing and clause boundaries. Pauses occur both within clauses and at clause boundaries without any strong cross-linguistic preference. However, the probability of a pause occurring increases as the syntactic boundary increases in strength. In Section 4.2, we found a gradient increase for most languages in the probability of a pause occurring within clauses to pauses at dependent clause boundaries and at main clause boundaries. This mirrors previous findings for English and French. For instance, Goldman-Eisler (1972: 105) found that temporal integration decreases as clause boundary independence increases in spontaneous spoken English, with 77.9% of all main clauses separated by pauses > 500ms, but less than a third of dependent clause boundaries. Grosjean & Deschamps (1972: 146) report that in their French data, over 60% of all silent pauses occur at main clause boundaries, with approximately 14% of the remaining pauses occurring at lower-level syntactic boundaries; Candea (2000: 166) similarly found that most of the pauses occur in her data at main clause boundaries with fewer pauses occurring at dependent clause boundaries and within clauses.

As our results are in line with previous findings, our study suggests that there is a crosslinguistically robust increase of the probability of pauses to occur within clauses, at dependent clause boundaries and at main clause boundaries. This finding is particularly remarkable given that the languages included in this study feature a great deal of variability in how dependent clauses are used and distributed.

Our second investigation in Section 4.3 found a two-way opposition in the duration of pauses occurring within clauses and those at clause boundaries. The attested lengths in this study match those found for Dalabon (Fletcher et al. 2004), and medium and long pauses in French (Campione & Véronis 2002). However, these comparisons should be taken with a grain of salt, as our methodological definition of pauses and locations, the choice of pause threshold, and the genres of texts involved all differ to different extents (see also Rochester 1973). Still, our results fit in with the general trend reported in the literature on pausing that silent pauses at clause boundaries tend to be longer than clauses that do not align with major syntactic boundaries (Candea 2000, Duez 1982, Goldman-Eisler 1972, Grosjean & Deschamps 1972).

In addition, we find a robust pattern across the languages surveyed here that there is no substantial difference between pause duration at main and dependent clause boundaries. This fits in with the findings reported by Candea (2000: 167), who only notes a significant difference in duration between pauses occurring at clause boundaries marking a discourse segments and pauses at other boundaries. This lack of difference between main and depen-

dent clause pause duration suggests that pauses do not simply reflect syntactic structures but help to structure discourse (at least in part) more generally.

6 Conclusion

Using naturalistic and typologically diverse spontaneous speech data, we found that clause boundaries were associated with pauses, with stronger boundaries more likely to co-occur with a pause. Pause duration showed a two-way distinction between pauses at clause boundaries and those which occurred within a clause. Furthermore, this study showed how existing grammatically annotated corpora can be used in the investigation of other research questions. With minimal extra annotation, we were able to conduct a first investigation into the role of silent pausing at the prosody-syntax interface in a range of languages not commonly represented in work on pause distribution and duration.

Acknowledgements

Glosses

ADV	adverb
COP	copula
DST	distal
DU	dual
EX	exclusive
EZ	ezafe
HABIT	habitual
IN	inclusive
INDF	indefinite
MOD	modifier
O	object
OBL	oblique
PL	plural
POSS	possessive
PREP	preposition
PRF	perfective
PST	past
RS	realis
SG	singular
TS	transitive suffix

References

Beattie, Geoffrey W. & B.L. Butterworth. 1979. Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech* 22(3). 201–211.

Belz, Malte & Jürgen Trouvain. 2019. Are 'silent' pauses always silent? In *19. international congress of phonetic sciences (ICPhS.*

Belz, Malte et al. 2017. Fluently disfluent?: Pauses and repairs of advanced learners and native speakers of German: *International Journal of Learner Corpus Research* 3(2). 118–148.

Bogomolova, Natalia, Dmitry Ganenkov & Nils N. Schiborr. 2021. Multi-CAST Tabasaran. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. Version 2108. Bamberg: University of Bamberg.

Bosker, Hans Rutger et al. 2013. What makes speech sound fluent? the contributions of pauses, speed and repairs. *Language Testing* 30(2). 159–175.

Brickell, Timothy. 2016. Multi-CAST Tondano. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. Version 2108. Bamberg: University of Bamberg.

Bürkner, Paul-Christian. 2017. Brms: an r package for bayesian multilevel models using stan. *Journal of Statistical Software* 80(1). 1–28.

Campione, Estelle & Jean Véronis. 2002. A Large-Scale Multilingual Study of Silent Pause Duration. In *Proceedings of Eurospeech 2002*, 199–202.

Candea, Maria. 2000. *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané. Etude sur un corpus de récits en classe de français.* Paris: Université Paris III.

Carpenter, Bob et al. 2017. Stan: a probabilistic programming language. *Journal of Statistical Software* 76(1). 1–32.

Chafe, Wallace L. 1994. *Discourse, consciousness, and time: the flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Christophe, Anne et al. 2008. Bootstrapping Lexical and Syntactic Acquisition. *Language and Speech* 51(1-2). 61–75.

Clark, Herbert H. & Jean E. Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition* 84(1). 73–111.

Cooper, William E. & Jeanne Paccia-Cooper. 1980. *Syntax and speech*. Cambridge, MA: Harvard University Press.

de Carvalho, Alex et al. 2017. Phrasal prosody constrains syntactic analysis in toddlers. *Cognition* 163. 67–79.

de Jong, Nivja H. 2016. Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching* 54(2). 113–132.

Duez, Danielle. 1982. Silent and non-silent pauses in three speech styles. *Language and Speech* 25(1). 11–28.

Ferreira, Fernanda. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30(2). 210–233.

Ferreira, Fernanda. 1993. Creation of prosody during sentence production. *Psychological Review* 100(2). 233–253.

Fletcher, Janet, Nicholas Evans & Belinda Ross. 2004. Pausing strategies and prosodic boundaries in Dalabon. In *Proceedings of the 10th Australian International Conference on Speech Science & Technology*, 436–439. Sydney: Macquarie University.

Frazier, Lyn, Katy Carlson & Charles Clifton. 2006. Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences* 10(6). 244–249.

Fuchs, Susanne et al. 2013. Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics* 41(1). 29–47.

Gee, James Paul & François Grosjean. 1983. Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology* 15(4). 411–458.

Goad, Heather, Natália Brambatti Guzzo & Lydia White. 2021. Parsing ambiguous relative clauses in l2 English: learner sensitivity to prosodic cues. *Studies in Second Language Acquisition* 43(1). 83–108.

Goh, Rui Zhe, Ian B. Phillips & Chaz Firestone. 2023. The perception of silence. *Proceedings of the National Academy of Sciences* 120(29). e2301463120.

Goldman-Eisler, Frieda. 1961a. A comparative study of two hesitation phenomena. *Language and Speech* 4(1). 18–26.

Goldman-Eisler, Frieda. 1961b. The Distribution of Pause Durations in Speech. *Language and Speech* 4(4). 232–237.

Goldman-Eisler, Frieda. 1968. *Psycholinguistics: Experiments in spontaneous speech*. London: The Academic Press.

Goldman-Eisler, Frieda. 1972. Pauses, clauses, sentences. *Language and Speech* 15(2). 103–113.

Grosjean, François & Maryann Collins. 1979. Breathing, pausing and reading. *Phonetica* 36(2). 98–114.

Grosjean, François & A. Deschamps. 1972. Analyse des variables temporelles du français spontané. *Phonetica* 26(3). 129–156.

Grosjean, François, Lysiane Grosjean & Harlan Lane. 1979. The patterns of silence: performance structures in sentence production. *Cognitive Psychology* 11(1). 58–81.

Haig, Geoffrey & Stefan Schnell. 2014. Annotations using GRAID (Grammatical Relations and Animacy in Discourse): Introduction and guidelines for annotators, Version 7.0.

Haig, Geoffrey & Stefan Schnell. 2021. Multi-CAST: Multilingual corpus of annotated spoken texts, Version 2108. Bamberg.

Haig, Geoffrey, Maria Vollmer & Hanna Thiele. 2015. Multi-CAST Northern Kurdish. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. Version 2108. Bamberg: University of Bamberg.

Hartsuiker, Robert J. & Lies Notebaert. 2010. Lexical Access Problems Lead to Disfluencies in Speech. *Experimental Psychology* 57(3). 169–177.

Hawthorne, Kara & LouAnn Gerken. 2014. From pauses to clauses: Prosody facilitates learning of syntactic constituency. *Cognition* 133(2). 420–428.

Kahng, Jimin. 2018. The effect of pause location on perceived fluency. *Applied Psycholinguistics* 39(3). 569–591.

Kendall, Tyler. 2013. *Speech Rate, Pause and Sociolinguistic Variation*. London: Palgrave Macmillan UK.

Kimoto, Yukinori. 2019. Multi-CAST Arta. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. Version 2108. Bamberg: University of Bamberg.

Kjelgaard, Margaret M. & Shari R. Speer. 1999. Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity. *Journal of Memory and Language* 40(2). 153–194.

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in english: acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59(5). 1208.

Krivokapić, Jelena. 2007. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics* 35(2). 162–179.

Krivokapić, Jelena, Will Styler & Benjamin Parrell. 2020. Pause postures: The relationship between articulation and cognitive processes during pauses. *Journal of Phonetics* 79. 100953.

Levelt, Willem J. M. 1993. *Speaking: From Intention to Articulation*. The MIT Press.

Maclay, Howard & Charles Osgood. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15(1). 19–44.

Mosel, Ulrike & Stefan Schnell. 2015. Multi-CAST Teop. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts*. Version 2108. Bamberg: University of Bamberg.

Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology: with a new foreword*. De Gruyter.

Petrone, Caterina et al. 2017. Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics* 61. 71–92.

R Core Team. 2021. *R: A language and environment for statistical computing*. manual. tex.organization: R Foundation for Statistical Computing. Vienna, Austria.

Rochester, S. R. 1973. The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research* 2(1). 51–81.

Sacks, Harvey, Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4). 696–735.

Schafer, Amy J. et al. 2000. Intonational Disambiguation in Sentence Production and Comprehension. *Journal of Psycholinguistic Research* 29(2). 169–182.

Schleef, Erik. 2019. The evaluation of unfilled pauses: Limits of the prestige, solidarity and dynamism dimensions. *Lingua* 228. 102707.

Seifart, Frank, Ludger Paschen & Matthew Stave (eds.). 2022. *Language Documentation Reference Corpus (DoReCo) 1.2.* Berlin & Lyon: Leibniz-Zentrum Allgemeine Sprachwissenschaft & laboratoire Dynamique Du Langage (UMR5596, CNRS & Université Lyon 2).

Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure.* Cambridge: MIT Press.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In *The handbook of phonological theory*, 2nd edn., 435–484. Oxford: Blackwell.

Shattuck-Hufnagel, Stefanie & Alice E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193–247.

Speer, Shari R. & Kiwako Ito. 2009. Prosody in first language acquisition - acquiring intonation as a tool to organize information in conversation: prosody in first language acquisition. *Language and Linguistics Compass* 3(1). 90–110.

Šturm, Pavel & Jan Volín. 2023. Occurrence and duration of pauses in relation to speech tempo and structural organization in two speech genres. *Languages* 8(1). 23.

Taboada, Maite. 2006. Spontaneous and non-spontaneous turn-taking. *Pragmatics* 16(2-3). 329–360.

Tannen, Deborah. 1985. Silence: Anything but. In Deborah Tannen & Muriel Saville-Troike (eds.), *Perspectives on silence*, 93–111. Norwood, NJ: Ablex.

Thieberger, Nick & Timothy Brickell. 2019. Multi-CAST Nafsan. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts.* Version 2108. Bamberg: University of Bamberg.

Truckenbrodt, Hubert. 1995. *Phonological phrases: Their relation to syntax, focus, and prominence.* Tübingen: University of Tübingen.

Truckenbrodt, Hubert. 2007. The syntax-phonology interface. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 435–456. Cambridge: Cambridge University Press.

Vollmer, Maria. 2020. Multi-CAST Mandarin. In Geoffrey Haig & Stefan Schnell (eds.), *Multi-CAST: Multilingual corpus of annotated spoken texts.* Version 2108. Bamberg: University of Bamberg.

Walker, Anne. 1985. The two faces of silence: The effect of witness hesitancy on lawyers' impressions. In Deborah Tannen & Muriel Saville-Troike (eds.), *Perspectives on silence*, 55–75. Norwood, NJ: Ablex.

Watson, Duane & Edward Gibson. 2004. The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19(6). 713–755.

Weilhammer, Karl & Susen Rabold. 2003. Durational aspects in turn taking. In *Proceedings of the International Conference of Phonetic Sciences*, 2145–2148. Barcelona.

Wennerstrom, Ann & Andrew F. Siegel. 2003. Keeping the floor in multiparty conversations: intonation, syntax, and pause. *Discourse Processes* 36(2). 77–107.

Yang, Li-chiung. 2004. Duration and pauses as cues to discourse boundaries in speech. In *Proceedings of speech prosody 2004*, 267–270. Nara, Japan.

Zellner, Brigitte. 1994. Pauses and the temporal structure of speech. In Eric Keller (ed.), *Fundamentals of speech synthesis and speech recognition*, 41–62. Chichester: John Wiley.