

Self Addressed Questions and Filled Pauses: A Cross-linguistic Investigation

**Ye Tian, Takehiko Maruyama &
Jonathan Ginzburg**

Journal of Psycholinguistic Research

ISSN 0090-6905

J Psycholinguist Res

DOI 10.1007/s10936-016-9468-5



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media New York. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Self Addressed Questions and Filled Pauses: A Cross-linguistic Investigation

Ye Tian¹  · Takehiko Maruyama² ·
Jonathan Ginzburg¹

© Springer Science+Business Media New York 2016

Abstract There is an ongoing debate whether phenomena of disfluency (such as filled pauses) are produced communicatively. Clark and Fox Tree (Cognition 84(1):73–111, 2002) propose that filled pauses are words, and that different forms signal different lengths of delay. This paper evaluates this Filler-As-Words hypothesis by analyzing the distribution of self-addressed-questions or SAQs (such as “what’s the word”) in relation to filled pauses. We found that SAQs address different problems in different languages (most frequently about memory-retrieval in English and Chinese, and about appropriateness in Japanese). In relation to filled pauses, British but not American English uses “um” to signal a more severe problem than “uh”. Chinese uses different filled pauses to signal the syntactic category of the problem constituent. Japanese uses different filled pauses to signal levels of interaction with the interlocuter. Overall, our data supports the Filler-As-Words hypothesis that filled pauses are used communicatively. However, the dimensions of its meanings vary across languages and dialects.

Keywords Self addressed questions · Filled pauses · Disfluency · Cross-linguistic analysis

Introduction

They had a new episode uh this past.. uh what is it.. Tue- Mon- Tuesday night maybe.
(Switchboard sw2560A-ms98)

We would like to acknowledge research funding from Agence Nationale de la Recherche, France.

✉ Ye Tian
tiany.03@gmail.com

¹ Laboratoire de linguistique formelle, UMR 7110, Université Paris Diderot, Sorbonne Paris Cité,
Paris, France

² Faculty of Letters, Senshu University, Kanagawa, Japan

The shooting is ...it's hard to describe, Coach Steve Kerr said, because I don't think we've ever seen anyone shoot the ball the way Steph does. (New York Times, May 24, 2015)

In natural conversations, disfluencies or phenomena of Own Communication Management (OCMs) are highly frequent,¹ on one estimate—found in about 6 out of 100 words (Fox Tree 1995). Reflected in the example above, speakers produce filled pauses (e.g. “uh”), make repairs (e.g. “Tue - Mon - Tuesday night”), and insert asides such as self-addressed questions (SAQs) (e.g. “what is it”). OCMs may reflect speakers' planning problems, or be used to communicate certain speaker intentions, such as the wish to hold a turn (Rochester 1973) or implying disagreement (Clark and Fox Tree 2002). This paper examines two OCM phenomena: filled pauses and self-addressed questions.

Previous Studies

Pauses in spontaneous speech was first studied from the perspective of language production, with the hypothesis that pauses (filled as well as silent) indicate problems in language generation (see a review by Rochester 1973). Some researchers (e.g. Goldman-Eisler 1968) propose that hesitation pauses after the first word of an utterance only reflects the transitional probability of each upcoming word. Others (e.g. Johnson 1965; Boomer 1965) argue that pauses reflect not only local but also distant planning, such as larger constituents (phrases, clauses, discourse units). Several studies considered whether filled pauses and silent pauses reflect different production processes. Maclay and Osgood (1959) proposed that pauses were filled when they are long, but this suggestion was not supported by a study by Boomer (1965). Levelt (1983) proposed that pauses and editing phrases in self-repair indicate that speakers have little access to their speech production process, but rather speakers manage their speech by self-monitoring based on parsing one's own inner or overt speech.

The interest in filled pauses later shifted from the perspective of language production to language comprehension. Some researchers claim that filled pauses are not noticed by listeners (e.g. Lickley and Bard 1996). However, the more popular and supported view is that filled pauses affect comprehension in certain ways. This stance, call it the “filler-as-word” hypothesis (FaW), has a weak and a strong version.

The weak version of FaW simply proposes that filled pauses have communicative functions, and they are different from silent pauses. The idea was first proposed by James (1972), who suggested that the English “uh” is an interjection signaling retrieval trouble. There is an abundance of evidence supporting the weak version of FaW. Siegman and Pope (1966) found that speakers produce more filled pauses in dialogues than in monologues. Pope et al. (1970) found that patients use longer silent pauses and fewer filled pauses when they are depressed than when they are anxious. Arnold et al. (2003) found that during reference identification, participants anticipate objects that are more difficult to name after a filled pause. Watanabe et al. (2008) found filled pauses cause listeners to expect complex constituents. Fraundorf and Watson (2011) found that participants remember a story better after listening to “Alice in Wonderland” with filled pauses than with coughs. Brennan and Williams (1995) showed that filled pauses differ from silent pauses in indicating the level of confidence of the speaker. Lake et al. (2011) showed that compared to the control group, speakers with Autistic Spectrum Disorders produce fewer filled pauses and more silent pauses. These studies suggest

¹ The term *own communication management* for what is also often called ‘self repair’ is due to Jens Allwood, see e.g., Allwood et al. (2005) for discussion.

that filled pauses have communicative values, but they did not address the issue of whether the communicative values differ for different filled pauses.

The strong version of FaW, proposed by [Clark and Fox Tree \(2002\)](#), states that not only do filled pauses in general have communicative import, but different pause fillers have different meanings. Clark and Fox Tree proposed that “uh” and “um” are conventional English words—they are interjections speakers use to announce a minor (uh) or a major (um) delay. Cross-linguistically, languages tend to have several pause fillers. Clark and Fox Tree postulated that languages need more than one pause filler to distinguish different lengths of delay. Compared to the weak version of FaW, evidence for the strong version is sparser. On the one hand, [Clark and Fox Tree \(2002\)](#) found in the London Lund corpus that “um” is more frequently followed by silent pauses than “uh”, and the pauses are longer. Similar findings have been reported by [Fox Tree \(2001\)](#). On the other hand, [O’Connell and Kowal \(2005\)](#) found no difference in the frequency or length of silent pauses after “um” and “uh” in media interviews of Hillary Clinton. [Brennan and Williams \(1995\)](#) found that when asked a question, answers preceded by “um” or “uh” do not give different impressions of the speaker’s confidence.

Our Research Questions

The studies above evaluated the strong FaW hypothesis mostly by looking at the length of pauses following different pause fillers. In this paper we take an alternative perspective: examining the relation between pause fillers and self addressed questions (SAQs). SAQs—a phenomenon of disfluency or OCM—are questions like “what’s the word” or “how to put it” that are inserted within utterances (often at an utterance-internal position)². Although we label them as being “self-addressed”, the addressee may also take the turn and respond. As far as we know, there has been no prior cross-linguistic study on SAQs, and little research on SAQs in general, with the exception of [Ginzburg et al. \(2014\)](#). Ginzburg et al. categorizes disfluencies into ones that are *backwards looking* or *forwards looking*. Backwards looking disfluencies are cases when an utterance is interrupted and replaced with an alteration that refers back to an already uttered reparandum. In this case, the speaker corrects or reformulates the reparandum. Forward looking disfluencies are cases when an utterance is interrupted by a filled or silent pause, but are continued without an alteration. SAQs only participates in forward looking disfluencies. They signal that the speaker is having planning problems with the upcoming utterance. This can be due to the fact that the speaker cannot remember certain information, cannot retrieve a word, or does not know how to appropriately phrase something. [Ginzburg et al. \(2014\)](#) propose that SAQs are a type of forward-looking disfluency. They show that SAQs have similarities with clarification questions. In English the contexts and contents of SAQs are regular: the most frequent context is where a noun phrase is expected; and the most frequent content is about memory retrieval, such as retrieving the names of people and places, remembering times and retrieving words.

How SAQs behave cross-linguistically is an interesting topic in itself, which, as far as we know, has not been previously studied. In relation to the FaW hypothesis, SAQs can provide potentially useful insights into the meanings of filled pauses for at least three reasons: SAQs and filled pauses often co-occur ([Ginzburg et al. 2014](#)); SAQs not only indicate that there IS a planning problem, as would also be indicated by a hesitation pause, but exactly what the planning problem is; SAQs should signal serious planning problems. We can assume this because inserting an SAQ gives the speaker more time to solve the planning problem than a filled/silent pause alone. So it is likely that planning problems addressed SAQs require

² In this paper we do not discuss non-disfluent utterances describing a state of self-oriented query, such as “I wonder when my mum will give me call”.

more time to solve. Also, an SAQ makes the planning problem explicit, which allows the possibility of help from the hearer. So, if different pause fillers are used to announce different length of delays, then the proposed long-pause fillers (e.g. “um”) should more frequently precede SAQs than short-pause fillers (e.g. “uh”).

In this study we ask two classes of questions:

Q1. The characteristics of SAQs: What problems do they address, what are their functions and distribution? Are these characteristics about SAQs cross-linguistically consistent? Based on results from [Ginzburg et al. \(2014\)](#), we hypothesize that the most frequent SAQs address memory retrieval problems, and we further hypothesize that the functions and distributions of SAQs are cross-linguistically consistent.

Q2. Evaluating the strong FaW hypothesis: Do pause fillers that supposedly signal longer delays also introduce more SAQs?

Based on predictions from [Clark and Fox Tree \(2002\)](#), we hypothesize that in English, SAQs are more frequently preceded by “um” than “uh”. In Chinese and Japanese, filled pauses that most frequently precede SAQs signal more serious planning problems.

To address these two questions, we investigated speech corpora in English (American and British), Mandarin Chinese and Japanese. We chose these three languages especially because of their different repertoires of filled pauses, but also because of their linguistic diversity. These three languages belong to different language families (Indo-European, Sino-Tibetan, and Japonic), and culturally diverse. Morphologically, English and Japanese are multi-syllabic languages while Mandarin Chinese is mono-syllabic. English and Japanese are inflected languages while Mandarin Chinese does not have any inflection ([Blevins 2006](#); [Bloch 1946](#); [Dai 1992](#)). English and Mandarin Chinese have an SVO word order, while Japanese has an SOV word order. However, while the word order in English is relatively rigid, it is much freer in Mandarin Chinese and Japanese ([Tamaoka et al. 2005](#); [Sun and Givón 1985](#)). As far as online processing and repair goes, previous studies ([Fox et al. 1996, 2009, 2010](#); [Hayashi 1994](#)) have shown that morpho-syntactic features influence the location and structure of self-repair. For example, [Fox et al. \(2010\)](#) found that the freer word order in German compared to English manifested in the fact that repairs in English frequently recycle back to the subject pronoun of the clause, it is much rarer in German. [Fox et al. \(2009\)](#) found that English speakers tend to initiate replacement repairs before the word is recognizably complete, while they tend to initiate recycling repairs after the word is recognizably complete. In comparison, Japanese speakers tend to initiate both types of repairs when the words are incomplete, while Mandarin Chinese speakers tend to initiate both types when the words are complete. It is possible that linguistic features such as word order and morphological complexity may also result in cross-linguistic differences in the distribution (especially the location) of filled pauses and SAQs.

In terms of the repertoires of filled pauses, in English, along with German, Dutch, Swedish, Norwegian, and Hebrew (incomplete list), the commonest filled pauses are “non-lexical”:³ in English, they are “um” and “uh” ([Clark and Fox Tree 2002](#)). On the other hand, in Chinese and Japanese, the commonest filled pauses include items derived from demonstratives: in Chinese “nage” (that) and “zhege” (this) ([Zhao and Jurafsky 2005](#)), and in Japanese “ano” (that) and “sono” (the) ([Watanabe et al. 2004](#))⁴. We will test the strong FaW hypothesis of Clark & Fox Tree in English, and see if similar results hold for languages with demonstrative filled pauses.

³ For a detailed discussion of the distribution of such hesitation markers in Germanic languages with variation across different populations, see [Wieling et al. \(2016\)](#).

⁴ Korean and Spanish also have demonstratives filled pauses ([Brody 1987](#); [Hayashi and Yoon 2010](#)).

We analysed data from five corpora, including Switchboard (Godfrey et al. 1992), the British National Corpus (Burnard 2000) and the London Lund corpus (Svartvik 1990), the NCCU Taiwan Mandarin corpus (Chui and Lai 2008) and the Corpus of Spontaneous Japanese (Maekawa 2003). In Sect. 2, we present the following analyses on each corpus⁵:

For Q1: Classification of SAQs and their frequencies; functions of SAQs (the types of problems addressed).

For Q2: Common pause fillers and their distributions; distributions of SAQs following different pause fillers.

In Sect. 3 we offer discussion of the results and some remaining open questions. Section 4 contains some brief conclusions.

Corpus Study

American English: Switchboard

SAQs and Their Distributions

The Switchboard corpus is a collection of about 2400 telephone conversations from speakers in all areas of the United States (Godfrey et al. 1992). To analyze SAQs in Switchboard, we first extracted a sample of 2500 utterances that contain a wh-question following a filled pause. From this sample, we extracted the most frequent forms of SAQ. We then searched for these frequent forms in the entire corpus, and manually checked whether the results contain SAQs. In total, we found a total of 365 regular SAQs in switchboard. Table 1 lists the SAQs and their frequencies.

Functionally, the majority of the SAQs found in this corpus address memory retrieval problems. Only one group—“how shall/ can I say/ put it”—addresses appropriateness/ phrasing problems. In order to see whether the function of the SAQs correlate with their syntactic contexts, we categorized the SAQs by their syntactic context, distinguishing whether the category of the anticipating constituent is a noun/NP, a clause/predicate, or others. For example, in the utterance “the motorway goes through, ah, what’s the name of the place? Wurley I think it is”, an NP (in this case a proper name) is expected at the point of the SAQ “what’s the name of the place”.

Here are some examples (SAQs are underlined):

- (1) a. Yeah well well that’s part of the how shall I say it that’s part of the experience I think. (sw4421A-ms98-a-0036)
- b. I know and it’s kind of um what’s the word I want I don’t it’s just, to me it’s just frightening you know. (sw2944B-ms98-a-0109)
- c. Oh I know especially if you get, what is it, Seclor I think that that just about breaks the bank right there. (sw2292B-ms98-a-0018)
- d. They’re they’re trying to get uh God what’s his name what’s that black man’s name uh. (sw3507B-ms98-a-0013)
- e. And uh when was it a couple weeks ago I was asked to go to uh jury duty. (sw2380B-ms98-a-0005)

⁵ These five corpora cover a time span of 18 years, which means that results suggesting cross-linguistic differences might potentially be due to language change. However, as the number of spontaneous speech corpora is limited, we could not find data from all target dialects and language matched in time period. Also, there are minor differences in the analyses across corpora due to the differences in annotation.

Table 1 SAQ forms, frequencies and contexts (in terms of the constituent anticipating)

SAQ	Occurrences	Anticipating					
		NP		Clause/predicate		Others	
What is/was it	130	123	95%	2	1.5%	5	3.7%
What do you/they call	74	69	93%	5	6.8%		
What is/was that_	29	29	100%				
What's his/her name	29	29	100%				
How shall/can I say/put it	27	4	15%	23	85%		
Who is it	17	17	100%				
What's the word	16	16	100%				
What's the name of	11	11	100%				
Where is it	11	11	100%				
What is it called	10	10	100%				
When is/was it	4	4	100%				
Which is it	3	3	100%				
What do I wanna say	2			2	100%		
Which one was that	2	2	100%				
Grand total	365	328	90%	32	8.8%	5	1.4%

In example 1a and 1b, the speakers are trying to find appropriate expressions, and used the SAQs “how shall I say it” and “what’s the word I want” respectively. In examples 1c and 1d, the speakers have difficulty retrieving proper names, and used the SAQs “what is it” and “what’s his name”. In example 1e, the speaker has difficulty remembering the time of an event, and used the SAQ “when was it”.

We can see that the vast majority of the SAQs (90%) appear in a position where the speaker is anticipating a noun or noun phrase. Only 8.8% of the SAQs (“how shall I say it”) appear where the speaker is anticipating a clause or a predicate. The functions of SAQs are highly correlated with their syntactic positions. Most of the SAQs (e.g. “what is it”, “what do you call...”) address the problem of memory retrieval, and they often appear where a noun phrase or a noun is expected. The other two SAQs, “how shall I say it” and “what do I wanna say”, signal that the speaker has a problem of phrasing something appropriately. These two often appear where a predicate or a clause is expected.

Fillers Preceding SAQs

We found that just under half (169 out of 365 or 46%) of the SAQs are preceded by a filler or a discourse marker: 115 (32%) by a filler and 54 (15%) by a discourse marker (“oh”, “well”, “you know”, “God/Gosh”, “yeah”, “hm”). Out of the fillers that precede SAQs, “uh” is much more frequent than “um”. Taking into account the different frequency of “uh” and “um” (86,528 “uh”s versus 37,319 “um”s in Switchboard), 0.11% of “uh” and 0.06% of “um” precede regular-form SAQs. Therefore, in Switchboard, if an SAQ is preceded by a pause filler, it is significantly more likely to be an “uh” than an “um” ($\chi^2 = 6.35$, $df = 1$ $p = .01$).

Overall, data from Switchboard (summarized in Table 2) show that speakers are more likely to use “uh” than “um” before an SAQ. This result does not support the hypothesis that “uh” introduces less severe planning problems than “um”.

Table 2 Fillers preceding SAQs in Switchboard

Filler	Preceding SAQs				Total frequency	Percentage
	NP	Clause/predicate	Other	Total		
uh	85	8	0	93	86,528	0.11
um	19	3	0	21	37,319	0.06

Table 3 SAQ distribution in the BNC

SAQ	Occurrences	Anticipating					
		Noun Phrase		Clause/predicate		Others	
What do you/they call	119	115	97%	4	3.4%		
What's his/her name	82	82	100%				
What is it	46	46	100%				
What's it called	28	28	100%				
How can/shall I say/ put/describe it	18	9	50%	6	33%	3	17%
What's the name of	17	16	94%	1	5.9%		
What's the word	17			6	35%		
Who is/was it	10	10	100%				
Where is/was it	9	9	100%				
When is/was it	8	8	100%				
Which one was it	1	1	100%				
Total	355	335	94%	14	3.9%	6	1.7%

British English: BNC

SAQs and Their Distributions

The British National Corpus (BNC) is a collection of written and spoken data of British English in the late twentieth century. The spoken part of the BNC has about 10 million words (Leech 1992). We found 355 regular-formed SAQs in the spoken part of the BNC (see Table 3). The exact forms and frequencies of the SAQs in BNC are different from Switchboard (the most frequent being “what do you/they call. . .” in BNC and “what is/was it” in Switchboard), however, like Switchboard, the vast majority (94%) of the SAQs appear where an NP is expected. The functions of the SAQs, as with Switchboard, correlate highly with their syntactic positions. Most SAQs that signal the problem of memory retrieval appear where an NP is expected. The SAQs about how to phrase something appropriately (“how can I say/put/describe it”) appear more often in the positions where a clause or a predicate is expected. For example, “*if people that er maybe erm, how can I put it, are not used t-taking everything back to a meeting and you know*” (BNC_HUXPS000). Note that these “appropriateness” SAQs can also appear where an NP is expected. For example: “*allowing all kinds of erm (pause) how can I put it foreigners in inverted commas, to vote in our elections in this country*” (BNC_JSG).

Table 4 Fillers preceding SAQs in the BNC

Filler	Preceding SAQs				Total frequency	Percentage
	NP	Clause/predicate	Other	Total		
er	60	1	1	62	89,948	0.07
erm	78	1	3	82	62,675	0.13

Fillers Preceding SAQs

As with Switchboard, in the BNC speech section, just under half of the SAQs (47.6%) are preceded by a pause filler or a discourse marker: 148 (42%) by a filled pause and 25 (7%) by a discourse marker (Table 4). However, the frequency pattern of preceding pause fillers are different from Switchboard. In Switchboard, the most frequent filler preceding SAQs is “uh”. In the BNC, on the other hand, it is “erm”. Taken into account the frequency of “erm” and “er” (62,675 vs. 89,948 occurrences), 0.13% of “erm” and 0.07% of “er” precede regular-form SAQs. Therefore, in the BNC, if an SAQ is preceded by a pause filler, it is significantly more likely to be an “erm” than an “er” ($\chi^2 = 14.37$, $df = 1$, $p = .0001$). The distribution of “um/erm” and “uh/er” preceding SAQs is different between the Switchboard and the BNC, signalled by a significant three-way (*corpus* by *um/uh* by *whether-preceding-SAQ*) interaction ($Z = -4.310$, $p = 1.63e-05$).

The pattern in the BNC supports the hypothesis that “erm” is used to announce a more serious planning problem than “er”.

British English: London-Lund Corpus

SAQs and Their Distributions

The London-Lund corpus is a collection of spontaneous conversations in British English containing about 500,000 words Svartvik (1990). We found a total of 37 SAQs. The sample is small, but the pattern resembles the distribution of SAQs in Switchboard and the BNC: most of the SAQs appear in an NP-anticipating position, and most SAQs address word retrieval problems (Table 5).

Fillers Preceding SAQs

Again, just under half of the SAQs (46%) were preceded by a filled pause (41%) or a discourse marker (5%).

The transcription of the filled pauses in London Lund distinguished prolonged from short forms, e.g. “u:m” and “um”. There are 1595 short “uh” and 1500 long “u:h” (in total 3095); 707 short “um”, and 1539 long “u:m” (in total 2246). Although “um/u:m” is a less frequent filled pause than “uh/u:h” (2246 vs. 3095), more SAQs were preceded by “um/u:m” than “uh/u:h” (10 vs. 5).

Data in the London Lund is small in size but the pattern reinforces the implication drawn from the BNC: “um” is used to announce a more serious planning problem than “uh” (Table 6).

Table 5 SAQ distribution in London Lund

SAQ	Occurrences	Anticipating	
		NP	Clause
What's it called	8	8	
What's his/her name	6	6	
What's the name of	5	5	
What do you call it	4	4	
What is/was it	4	3	1
What did I/was I going to say	4	2	2
What's the word	2	2	
How can I put it	1	1	
Other	3		
Total	37	34	3

Table 6 Fillers preceding SAQs in London Lund

Filler	Preceding SAQs				Total frequency	Percentage
	NP	Clause/predicate	Other	Total		
u:m	6	1	0	7	1539	0.45
um	3	0	0	3	707	0.42
uh	3	1	0	4	1595	0.25
u:h	0	1	0	1	3095	0.03

Chinese: NCCU Taiwan Mandarin Corpus

SAQs and Their Distributions

The NCCU Corpus of Spontaneous Chinese (Chui and Lai 2008) contains face-to-face conversations (not necessarily between just two speakers) in three languages: Mandarin, Hakka, and Southern Min. The Mandarin sub-corpus contains about 3.5 h of conversations, 4459 turns and 105,481 characters.

We found 48 SAQs (Table 7). They appeared in the form of “what” (shenme), “who” (shui) and “where” (na/nali), which address memory retrieval problems, and “how to say it” (zenme jiang), which address the problem of phrasing appropriateness. Like English, most

Table 7 SAQ distribution in NCCU Mandarin

SAQ	Occurrences	Anticipating			
		Noun Phrase	Clause/predicate		
What (什么 shen me)	24	20	83%	4	17%
Who (谁 shui)	12	12	100%		
How to say it (怎么讲 zen me jiang)	10			10	100%
Where (哪 na)	2	2	100%		
Total	48	34		14	

of the SAQs appear in an NP-anticipating position. Most SAQs address memory retrieval problems.

Here are some examples. SAQs are underlined (“.” signals silent pauses).

- (2) a. 然后还有那个...什么...手震..它有防手震

Then there is nage... what... shaking..it has anti-shaking.

- b. 然后我们就讲到...谁..eh..好像是在讲你跟小黄鱼..

Then we were talking about ... who... eh.. I think (we) were talking about you and Xiaohuangyu..

- c. 喔对了你不是坐里面哪...南海特区

Oh right weren't you sitting inside where... the South Sea District)

- d. 可是全科班有时候好像..就是...怎么讲呢以前都是要划位的...

But it seems the whole class sometimes ...how to say it, were allocated seats in the past.

In example 2a, the speaker had difficulty retrieving the word “shou zhen” (shaking movements by hands), and used the SAQ “shen me” (what). In example 2b, the speaker could not remember the name of a person, and used the SAQ “shui” (who). In example 2c, the speaker could not remember the name of a location, and used the SAQ “na” (where). In example 2d, the speaker was trying to find an appropriate expression, and used the SAQ “zenme jiang ne” (how to say it).

Fillers Preceding SAQs

In Mandarin, the demonstratives “nage” (that) and “zhege” (this) can be used as pause fillers. In addition, there are “non-lexical” fillers similar to “um” and “uh”, such as “en” and “eh”. The distributions of different pause fillers preceding SAQs are summarized in Table 8.

About half (54%) of the SAQs are preceded by a pause filler. The most frequent pause filler preceding a SAQ is “nage”. Does this suggest that “nage” is used to signal longer delays than other pause fillers in Chinese? No. Firstly, the percentage of “nage” followed by a silent pause is not significantly different from any other fillers ($\chi^2 = 4.7, p = .19$), and the length of pause following “nage” is also comparable to that following other fillers ($t_s < 1.5, p_s > 0.11$). Secondly, a careful look at the distribution of SAQs reveals that the choice of pause filler before a SAQ is driven mainly by the syntactic context. All but one SAQs followed by “nage” occurred in contexts where a noun phrase is expected (e.g. “shen me”, meaning “what”). On the other hand, among the 15 SAQs in clause/ predicate anticipating contexts (e.g. how to say it), only 1 was preceded by “nage”. The rest were preceded by “um”, “eh”, the editing phrase “jiushi” (that-is), or a silent pause.

The results from the Mandarin Chinese corpus are consistent with the broad hypothesis that different pause fillers are used to announce different types of planning problems. However the “types” here are not levels of severity, but the syntactic categories of the problem constituent.

Table 8 Fillers preceding SAQs in NCCU Mandarin

Filler	Preceding SAQs				Total frequency	Percentage
	NP	Clause/predicate	Other	Total		
nage	20	1	0	21	392	5.40
en/um	0	3	0	3	121	2.50
eh/ah	2	0	0	2	258	0.80
zhege	0	0	0	0	24	0

Table 9 SAQ distribution in CSJ

SAQ	Occurrences	Anticipating		Clause	
		Noun phrase			
How to say it (nante yu: ndesuka ne)	52	7	13.5%	45	86.5%
What is it (nan daro)	33	12	36.4%	21	63.6%
What was that again (nan dakke)	14	12	85.7%	2	14.3%
Total	99	31		68	

Japanese: Corpus of Spontaneous Japanese

SAQs and Their Distributions

The Corpus of Spontaneous Japanese (CSJ) is a large-scale and richly annotated spontaneous speech corpus of standard Japanese1 (Maekawa 2003). It consists of approximately 661 hours of spontaneous speech including 7.52 million words, collected from 3302 speeches by 1417 different speakers. Most of the recorded materials are spontaneous monologues, while it also includes a small data set of dialogue and read speech.

In this study we use 58 dialogues with 153,591 words, approximately 12.2 h. We found a total of 99 SAQs (Table 9). These can be classified into three groups, “what was that again?” with six grammatical variations (e.g. nan dakke, nan deshitakke), “what is it?” with six variations (e.g. nan daro:, nan desho:), and “how to say it” with 19 variations (e.g. nante yu: ndesuka ne, nante yu: ndaro:).

The result (Table 9) shows a clear difference among the three groups. The SAQs of “what was that again” are likely to be used when the speaker tries to remember a noun phrase, and they appear mostly in NP-anticipating contexts. On the other hand, the SAQs of “how to say it” frequently occur when the speaker tries to produce a new clause to pursue their explanation, and they mostly appear in clause-anticipating contexts. The SAQs of “what is it” are used for the both cases.

Examples of each type are shown below. “(741ms)” represents a pause of 741 ms (only pauses over 50ms were transcribed), and (F**) represents filled pauses. SAQs are underlined.

- (3) a. (F *ma*) *ii imi de nante yu: ndesuka ne* (F *ano:*) (273ms) *tatakiage tteyu: koto desu ne*
 well, in a good sense, how shall I say, ahm, (273ms) he worked his way up from the bottom.
- b. (F *e:to*) *modan dansu tte yu: no wa nandaro:* (85ms) *kurasikku baree to mo chotto chigau shi*
 well, what I call modern dance is, what is it, (85ms) it's a bit different from the classic ballet, for one thing,
- c. *hoteru no choshoku toka mo* (741ms) *nanka nan dakke tomato marugoto ikko yaita yatsu toka*
 the breakfast in the hotel for example was, (741ms) like, what was that again, a whole baked tomato, for example,

In example 3a, the speaker tries to remember the breakfast she had in the hotel, and uses the SAQ “*nan dakke*” while she’s thinking. Then she remembers it, and utters the noun phrase. On the other hand, the speaker in example 3b, explains what modern dance is, and uses the SAQ “*nandaro:*” in the middle of her utterance. At this moment she tries to find a good explanation, and then she continues her speech with a new clause. 3c is an example where the speaker tries to explain how his friend worked hard, and uses the SAQ “*nante yu: ndesuka ne*”. In this case he buys time to continue his speech, eventually producing a new clause to explain what he had in mind.

Unlike English and Chinese, in the CSJ, more SAQs appear in positions anticipating clauses rather than NPs (about 2:1). While most SAQs address memory retrieval problems in English and Chinese corpora, in the CSJ, these SAQs (*nan dakke*, *nan daro*) are less frequent, accounting for 47% of all SAQs. The most frequent SAQ “how shall I say it” (*nante yu: ndesukane*) is used when the speaker tries to phrase something appropriately.

Fillers Preceding SAQs

As with Chinese, in Japanese, pause fillers include both “non-lexical” forms (e.g. *e*, *n*) and lexical fillers that can be used as demonstratives (e.g. *ano* (that), *sono* (it)). Table 10 summarizes the distribution of SAQs preceded by different pause fillers and the total frequency of these pause fillers. Note that these figures come from searches in the entire CSJ, rather than our sample described in Sect. 2.5. CSJ annotates elongation of words with “:”. The fillers in Table 10 include their elongated variations, but most fillers preceding SAQs are elongated.

We can see that the most frequent fillers preceding SAQs are “*ano*”, followed by “*sono*”, yet the most frequent pause filler in general is “*e*” (twice as frequent as “*ano*” and “*sono*” combined). Why are “*ano*” and “*sono*” the preferred pause fillers for introducing SAQs? Is it because “*ano*” and “*sono*” are used in different syntactic positions from “*e*” (like Chinese)? Is it because “*ano*” and “*sono*” signal more severe problems than “*e*”? Is there any other explanation?

To evaluate the first possibility, we compared the frequencies of different pause fillers preceding SAQs in different syntactic positions. About half of the SAQs in our sample are preceded by a pause filler (51 out of 99 or 52%). As “*ano:*” and “*sono:*” are derived from demonstratives, it is possible that, as in the case for Chinese, they more frequently precede NP-anticipating SAQs. However, this is not the case. Across all pause fillers, they more often precede clause-anticipating than NP-anticipating SAQs (see Table 11). This suggests that Japanese demonstrative pause fillers do not function like Chinese demonstrative pause fillers. They do not signal the syntactic category of the problem constituent.

Table 10 Fillers preceding SAQ in CSJ

Pause filler	Preceding SAQs	Total frequency	Percentage
ano:/a:no:/ano	255 (206 elongated)	79037	0.32
sono:/sono	119 (63 elongated)	23357	0.51
e:/e	115 (101 elongated)	168747	0.07
n:/n	62 (38 elongated)	14756	0.52
ma:	57	32620	0.17
a:	21	10860	0.19
o:	16	8613	0.19
u:	11	3398	0.32

To evaluate the second possibility, we examined the length of silent pauses following different pause fillers. If “ano:” and “sono:” signal more serious problem than “e”, then “ano:” and “sono:” should be followed by longer pauses than “e”. We could not find any literature directly comparing pauses following different pause fillers in Japanese. A related study by [Quimbo et al. \(1998\)](#) analyzed the prosodic features of fillers and the same form uttered as regular words. They found that when speakers are asked to read sentences containing pause fillers, the average pause length after “ano” was 0.3 s. They did not report the mean pause duration after “e”, but plotted individual data points in a figure ([Quimbo et al. 1998](#), figure 5). It can be estimated that the mean pause length following read “e” is between 0.3 and 0.4 s, similar to that of “ano”. Thus the hypothesis that “ano” and “e” are used to signal different lengths of delays is not supported.

What, then, could be the explanation? [Sadanobu and Takubo \(1995\)](#) propose that “ano” indicates that the speaker is trying to choose the suitable expressions for the current context, and thus gives the impression of politeness (see also [Iwasaki 2011](#)). On the other hand, the filler “eto” (sometimes grouped with “e” as “non-semantic fillers” cf. [Kawamori et al. \(1996\)](#)), is, arguably, used to distance oneself from the interlocuter in order to enable one’s own reflection, analogous to avoiding eye contact due to cognitive difficulty ([Doherty-Sneddon and Phelps 2005](#)). This is consistent with the fact that “eto”/“e” but not “ano” can be used when the speaker is alone ([Sadanobu and Takubo 1995](#)). Our data also suggest that “ano”/“sono” on the one hand and “e” on the other hand may have different pragmatic functions. In general, speakers use “e” more frequently because most of the planning problems do not need to be made public. However, SAQs are acts that invite rather than shut off interaction from the interlocuter. Thus when a planning problem is made public by an SAQ, the speaker is more likely to use a filler that signals involvement of the interlocuter, which could also explain why “ano” shows politeness.

Overall, we see that in the CSJ, the most frequent SAQs are used when the speaker is attempting to phrase something appropriately, rather than having a problem with memory retrieval. More SAQs appear in clause-anticipating positions than NP-anticipating positions. In terms of filled pauses preceding SAQs, although “e” is the most frequent filled pause overall, when preceding an SAQ, it is less frequent than the demonstrative filler “ano”.

Our analysis in CSJ suggests yet another dimension of the meaning of pause fillers: they can have a pragmatic function, signaling how much the speaker is inviting the interlocuter to interact regarding the present question under discussion.

Table 11 Number of fillers preceding different types of SAQs in CSJ

Pause filler preceding SAQ	Anticipating	
	NP	Clause
ano:/a.no:/ano	6	9
e:/e	2	7
e:to/eto	2	4
sono:/sono	0	5
ma:	0	4
a:	1	2
ko:	1	2
u:n	1	2
n:	0	2
u:	0	1
Total	13	38

Discussion

Our study addresses two topics: the distribution of SAQs and whether they support the Fillers-as-Words hypothesis. In this section, we discuss the findings for each question separately.

SAQ

SAQs are questions inserted mid-utterance when the speaker is encountering a planning problem. We analysed the forms and contexts of SAQs in five corpora. There are findings that apply to all languages considered, as well as language-specific features of SAQs. For a start, we found that SAQs exist in all the corpora we surveyed; these languages all allow questions to be inserted mid-utterance. The hearers usually have no problem incrementally processing such deviations, and can incorporate the answers to SAQs into the original utterance, or complete the utterance collaboratively by providing an answer. We hypothesize that the existence of SAQs is a linguistically universal phenomenon.

A second finding applying all our corpora is that about half of the SAQs are “bare”: they are not preceded by any filled pauses or discourse markers. Some of the “bare” SAQs are preceded by silent pauses, but not all. If filled pauses are words used to announce a planning problem, such announcement is not obligatory.

Apart from the two quasi-universal findings above, there are language/dialect specific features. Firstly, the frequency of SAQs in general varies. Per 100,000 words, there are 11 SAQs in Switchboard, 4 in the British National Corpus, 7 in London Lund, 46 in the NCCU Taiwan Mandarin corpus, and 64 in the Corpus of Spontaneous Japanese. It seems SAQs are more frequent in the Chinese and Japanese data than in the English data. The different corpus settings could certainly play a role in the frequency of SAQs: conversations in Switchboard are carried out over the phone by strangers, while conversations in the BNC, London Lund and NCCU Mandarin are by acquaintances/friends. The majority of the data in CSJ Japanese are monologues, but the sample analysed for SAQ contains only dialogues. It is possible that these features can cause differences in SAQ frequency, but they cannot readily explain on the one hand the lack of differences among the (American/British) English corpora and also the much higher frequencies in Chinese and Japanese. So it is possible that SAQs are used much more frequently in some languages than others.

Secondly, the functions of SAQs differ between our data in English and Chinese, and the corresponding Japanese data. British English, American English and Taiwan Mandarin reveal that the majority of SAQs address the problem of memory retrieval, and they appear more frequently in an NP-anticipating position than a clause/predicate anticipating position. On the other hand, in Japanese, the most frequent SAQs address the problem of phrasing something appropriately, and more often than not they appear in a predicate/ clause-anticipating than an NP-anticipating positions. One might conjecture that this difference has a syntactic cause, perhaps in Japanese: predicates have more significance than noun phrases. Compared to English, Japanese allows pro-drop, and therefore null-subject phenomena (utterances without subjects) are more frequent in Japanese than in English (Zushi 2003). However, null-subjects are also frequent in Chinese (Wang et al. 1992), therefore this cannot be an explanation for the more frequent predicate/clause-anticipating SAQs in Japanese. Therefore we cautiously conclude that Japanese speakers produce more appropriateness related SAQs than speakers of Chinese or English. In future studies we will test this conclusion using different genres of speech corpora.

These findings speak against our first hypothesis that the functions and distributions of SAQs are cross-linguistically consistent. Rather, both the frequency and the functional distributions of SAQs vary cross-linguistically.

SAQs in Relation to Filled Pauses

The second goal of this study is to evaluate the Fillers-as-Words hypothesis (FaW). The weak version of FaW states that filled pauses have different functions from silent pauses, and the strong version of FaW proposes that different filled pauses have different meanings. Specifically, Clark and Fox Tree (2002) propose that languages use different pause fillers to announce problems of different severities. In English, “um” is used to announce a more serious problem than “uh”, and they found that “um” is followed by longer delays than “uh”. In our study, we reasoned that since SAQs signal relatively serious planning problems. Based on the prediction of Clark and Fox Tree (2002), we hypothesized that in English more SAQs should be preceded by “um” than “uh”; in Chinese and Japanese, pause fillers that are followed by longer silences should precede more SAQs. This hypothesis was not fully supported by our results.

For English, our data from the BNC and London Lund supports the strong FaW: SAQs are more likely to be preceded by “um” than “uh”. However, the pattern is reversed in Switchboard. The results suggests that there is a dialectal difference between British and American English in the meaning of “uh” and “um”. This implication is consistent with the findings of O’Connell and Kowal (2005) and Brennan and Williams (1995). O’Connell and Kowal (2005) found no difference in silence following “uh” and “um” in interviews of Hillary Clinton. Brennan and Williams (1995) found (using American participants) that answers preceded by “um” and “uh” do not give different impressions of the speaker’s confidence, and their study was one in American English with American subjects.

In Chinese, we found that SAQs are more likely to be preceded by the demonstrative filler “nage” than other fillers. In general, “nage” does not precede longer pauses than other fillers. Why is “nage” favoured by SAQs? We found that SAQs in an NP-anticipating position are more likely to be preceded by “nage” than SAQs in a clause/predicate position (see Zhao and Jurafsky (2005) for a similar finding). This suggests that in Chinese, different pause fillers do not announce different severities of the planning problem, but signal the syntactic category of the problem constituent.

For Japanese, we found that pause filler “e” is more frequent than fillers of demonstrative origin such as “ano” and “sono”. However, before an SAQ, “ano” and “sono” are more frequent than “e”. Unlike “nage” in Chinese, “ano” and “sono” do not more frequently precede NP-anticipating SAQs than clause-anticipating SAQs. Nor is there evidence for “ano” announcing longer pauses than “e”. This result points to yet another dimension of the meaning of pause fillers: signaling levels of interaction with the interlocuter. “Ano” and “sono” are used to invite interaction from the interlocuter (which gives the impression of politeness), while “e” is used to signal a need for more private cognitive processes. This idea is in line with the finding of [Watanabe et al. \(2006\)](#) that “e” is used more frequently in academic lectures while “ano” is used more frequently in casual talk.

Our data supports the strong FaW hypothesis that languages use different fillers to communicate different meanings. However, it refines the proposal of [Clark and Fox Tree \(2002\)](#) in that we found that the meaning of fillers can have different dimensions in different languages. It may announce different delays (British English), but it may also signal the syntactic category of the problem constituent (Mandarin Chinese) or signal the level of interaction with the interlocuter (Japanese).

Conclusion

Our study evaluated the distribution of SAQs and its relation with filled pauses in corpora of American English, British English, Chinese and Japanese. We found that in all corpora, SAQs exist and about half of them are preceded by filled pauses (and occasionally by discourse markers). The frequency of SAQs is lower in English than in Chinese and Japanese. In English and Chinese, the most frequent SAQs address memory retrieval problems, and appear in an NP anticipating positions. In Japanese, the most frequent SAQs address appropriateness problems, and appear in clause anticipating positions. In relation to filled pauses, British but not American English uses “um” to signal a more severe planning problem than “uh”. Chinese uses different pause fillers to signal the syntactic category of the problem constituent. Japanese uses different pause fillers to signal levels of interaction with the interlocuter.

Overall, we propose that SAQs are a universal phenomenon, as consistent with the analysis of forward looking OCMs by [Ginzburg et al. \(2014\)](#). However, their usage (frequency, problems addressed) varies across languages. Filled pauses are consciously produced and are communicative, but the dimensions of meanings of filled pauses vary across languages and dialects.

Funding This study was funded by the French Investissements d’Avenir–Labex EFL program (ANR-10-LABX-0083) and by the Disfluences, Exclamations, and Laughter in Dialogue (DUEL) project within the Projets Franco-Allemand en sciences humaines et sociales of the Agence Nationale de Recherche (ANR) and the Deutsche Forschungsgemeinschaft (DFG), Grant Number ANR-13-FRAL-0001-01. Ginzburg is also supported by a senior member fellowship from the Institut Universitaire de France.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Allwood, J., Ahlsén, E., Lund, J., & Sundqvist, J. (2005). Multimodality in own communication management. In *Proceedings of the second Nordic Workshop on Multimodal Communication*, Gothenburg.

- Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, 32(1), 25–36.
- Blevins, J. P. (2006). 22 English inflection and derivation. *The handbook of English linguistics*.
- Bloch, B. (1946). Studies in colloquial Japanese I inflection. *Journal of the American Oriental Society*, 66(2), 97–109.
- Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech*, 8(3), 148–158.
- Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, 34(3), 383–398.
- Brody, J. (1987). Particles borrowed from Spanish as discourse markers in Mayan languages. *Anthropological Linguistics*, 29, 507–521.
- Burnard, L. (2000). *Reference guide for the British National Corpus (World Edition)*. Oxford: Oxford University Computing Services.
- Chui, K., Lai, H.-L., et al. (2008). The nccu corpus of spoken chinese: Mandarin, hakka, and southern min. *Taiwan Journal of Linguistics*, 6, 119–144.
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84(1), 73–111.
- Dai, X.-L. (1992). *Chinese morphology and its interface with syntax*. PhD thesis, The Ohio State University.
- Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & Cognition*, 33(4), 727–733.
- Fox, B., Wouk, F., Hayashi, M., Fincke, S., Tao, L., Sorjonen, M.-L., et al. (2009). A cross-linguistic investigation of the site of initiation in same-turn self-repair. In *Conversation analysis: Comparative perspectives* (pp. 60–103) Cambridge, UK: Cambridge University Press.
- Fox, B. A., Hayashi, M., & Jaspersen, R. (1996). Resources and repair: A cross-linguistic study of syntax and repair. *Studies in Interactional Sociolinguistics*, 13, 185–237.
- Fox, B. A., Maschler, Y., & Uhmman, S. (2010). A cross-linguistic study of self-repair: Evidence from english, german, and hebrew. *Journal of Pragmatics*, 42(9), 2487–2505.
- Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, 34(6), 709–738.
- Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, 29(2), 320–326.
- Fraundorf, S. H., & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language*, 65(2), 161–175.
- Ginzburg, J., Fernández, R., & Schlangen, D. (2014). Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*, 7(9), 64.
- Godfrey, J. J., Holliman, E. C., & McDaniel, J. (1992). Switchboard: Telephone speech corpus for research and development. In *Proceeding of 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 1, pp. 517–520. IEEE.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- Hayashi, M. (1994). A comparative study of self-repair in English and Japanese conversation. *Japanese/Korean Linguistics*, 4, 77–93.
- Hayashi, M., & Yoon, K.-E. (2010). A cross-linguistic exploration of demonstratives in interaction. *Fillers, Pauses and Placeholders*, 93, 33.
- Iwasaki, N. (2011). Filling social space with fillers: Gains in social dimension after studying abroad in Japan. *Japanese Language and Literature*, 45, 169–193.
- James, D. (1972). Some aspects of the syntax and semantics of interjections. In *Eighth regional meeting of the Chicago Linguistic Society*, pp. 162–172.
- Johnson, N. F. (1965). The psychological reality of phrase-structure rules. *Journal of Verbal Learning and Verbal Behavior*, 4(6), 469–475.
- Kawamori, A., Kawabata, M., & SHIMAZU, T. (1996). A phonological study on Japanese discourse markers. *Language*, 297, 306.
- Lake, J. K., Humphreys, K. R., & Cardy, S. (2011). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. *Psychonomic Bulletin & Review*, 18(1), 135–140.
- Leech, G. (1992). 100 Million words of English: The British National Corpus (BNC). *Language Research*, 28(1), 1–13.
- Levitt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, 14(1), 41–104.
- Lickley, R. J., & Bard, E. G. (1996). On not recognizing disfluencies in dialogue. In *Spoken Language, 1996. Proceedings of Fourth International Conference on ICSLP 96*, Vol. 3, pp. 1876–1879. IEEE.
- Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous english speech. *Word*, 15(1), 19–44.

- Maekawa, K. (2003). Corpus of spontaneous Japanese: Its design and evaluation. In *ISCA & IEEE workshop on spontaneous speech processing and recognition*.
- O'Connell, D. C., & Kowal, S. (2005). Uh and um revisited: Are they interjections for signaling delay? *Journal of Psycholinguistic Research*, 34(6), 555–576.
- Pope, B., Blass, T., Siegman, A. W., & Rahe, J. (1970). Anxiety and depression in speech. *Journal of Consulting and Clinical Psychology*, 35(1p1), 128.
- Quimbo, F. C., Kawahara, T., & Doshita, S. (1998). Prosodic analysis of fillers and self-repair in Japanese speech, pp. 3313–3316.
- Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2(1), 51–81.
- Sadanobu, T., & Takubo, Y. (1995). The monitoring devices of mental operations in discourse: A case of 'eeto' and 'ano(o)'. *Gengo Kenkyu*, 108, 74–93.
- Siegman, A. W., & Pope, B. (1966). Ambiguity and verbal fluency in the tat. *Journal of Consulting Psychology*, 30(3), 239.
- Sun, C.-F., & Givón, T. (1985). On the so-called sov word order in mandarin Chinese: A quantified text study and its implications. *Language*, pp. 329–351.
- Svartvik, J. (1990). *The London-Lund corpus of spoken English: Description and research* (Vol. 82). Lund: Lund University Press.
- Tamaoka, K., Sakai, H., Kawahara, J.-I., Miyaoka, Y., Lim, H., & Koizumi, M. (2005). Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research*, 34(3), 281–332.
- Wang, Q., Lillo-Martin, D., Best, C. T., & Levitt, A. (1992). Null subject versus null object: Some evidence from the acquisition of Chinese and English. *Language Acquisition*, 2(3), 221–254.
- Watanabe, M., Den, Y., Hirose, K., & Minematsu, N. (2004). Clause types and filled pauses in Japanese spontaneous monologues. In *INTERSPEECH*.
- Watanabe, M., Den, Y., Hirose, K., Miwa, S., & Minematsu, N. (2006). Factors affecting speakers' choice of fillers in Japanese presentations. In *Proceedings of Interspeech*, 2006, pp. 1256–1259.
- Watanabe, M., Hirose, K., Den, Y., & Minematsu, N. (2008). Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication*, 50(2), 81–94.
- Wieling, M., Grieve, J., Bouma, G., Fruehwald, J., Coleman, J., & Liberman, M. (2016). Variation and change in the use of hesitation markers in germanic languages. *Language Dynamics and Change*.
- Zhao, Y., & Jurafsky, D. (2005). A preliminary study of mandarin filled pauses. In *Proceedings of DiSS*, 2005, pp. 179–182.
- Zushi, M. (2003). Null arguments: The case of Japanese and romance. *Lingua*, 113(4), 559–604.