Defining Filler Particles: A Phonetic Account of the Terminology, Form, and Grammatical Classification of "Filled Pauses"

Malte Belz (ID)

Department of German Studies and Linguistics, Humboldt-Universität zu Berlin, 10099 Berlin, Germany; malte.belz@hu-berlin.de

Abstract: The terms *hesitation, planner, filler*, and *filled pause* do not always refer to the same phonetic entities. This terminological conundrum is approached by investigating the observational, explanatory, and descriptive inadequacies of the terms in use. Concomitantly, the term *filler particle* is motivated and a definition is proposed that identifies its phonetic exponents and describes them within the linguistic category of particles. The definition of filler particles proposed here is grounded both theoretically and empirically and then applied to a corpus of spontaneous dialogues with 32 speakers of German, showing that in addition to the prototypical phonetic forms, there is a substantial amount of non-prototypical forms, i.e., 9.5%, comprising both glottal (e.g., [ʔ]) and vocal forms (e.g., [ɛɸ], [jɛvə]). The grammatical classification and the results regarding the phonetic forms are discussed with respect to their theoretical relevance in filler particle research and corpus studies. The phonetic approach taken here further suggests a continuum of phonetic forms of filler particles, ranging from singleton segments to multi-syllabic entities.

Keywords: filler; filled pause; hesitation; filler particle; phonetic form; definition; corpus study; interjection; spontaneous speech; continuum

1. Introduction

This study sets out to achieve two goals. The first goal is to present arguments in favour of the term *filler particle* (for a full definition, see Section 3) instead of other terms, e.g., *filled pause* or *hesitation*. Although filler particles (e.g., /ɛː ɛːm m/ in German and /ə əːm m/ in English) have been studied extensively since the 1950s (Maclay and Osgood 1959), the terminological variation is substantial (e.g., *hesitation, hesitation marker, filled pause, filler, disfluency, disfluency marker, pause-filler, hesitator, planner, discourse particle, discourse marker*). This variation is at least in part a consequence of theoretical viewpoints influencing the choice of term to highlight certain functional aspects (e.g., *planner* is used by Jucker 2015 and Tottie 2016 to highlight the planning function). The terminological review undertaken here aims to funnel the current understanding of the phenomenon as evidenced by the literature back into the referent describing it, using a term that fulfils the requirements of being precise, unambiguous, non-judgemental (Gläser 1995, p. 528), and without an a priori determined pragmatic function. The theoretical viewpoint of this study focuses on the grammatical part-of-speech status, resulting in the choice of *filler particle*. In order to undergird this novel term, the legacy terms *hesitation, filler*, and *filled pause* will be discussed regarding the dimensions of form, function, and linguistic category.

The second goal of this study is to provide a definition of filler particle which is theoretically grounded in usage-based grammatical theory. Drawing from the literature, I argue that the phenomenon should be categorized as a *particle* in the linguistic part-of-speech sense. Importantly, the definition is also intended to be applicable for the identification of unseen phonetic exponents of filler particles in the acoustic material.

In this respect, the study is clearly phonetically oriented as well as corpus-based, but it may serve as a foundation for studies in conversational and interactional linguistics.

Both of these research goals are supported in the following two sections using corpus-based data from spontaneous German dialogues. The definitory concept for the phonetic exponents of filler particles (Section 3) is then evaluated by means of a corpus-based study. It will be shown that this open-set approach substantially extends the range of phonetic forms beyond the prototypical vocalic and vocalic–nasal forms and is necessary to paint and understand the full picture. The data and their annotation are presented in Section 4. The frequencies and forms of filler particles are presented in Section 5, followed by a discussion in Section 6.

2. Terminological Difficulties

The aim of this section is to briefly sketch the difficulties with existing terms for filler particles, and, at the same time, to provide argumentative support for analysing filler particles as particles in a grammatical sense. To be as clear as possible, some terminological housekeeping is in order. I will use the term *filler particle* to refer to the phonetic exponents subsumed by the definition in Section 3. The acoustic representation of filler particles is referred to by the term *phonetic exponents* (Kohler et al. 2005), understood as an identifiable and delimitable interval which is sufficiently dissimilar to surrounding speech segments.[1] Finally, to remain consistent, *filler particle* is also used when the cited studies use other terms.

2.1. Form—Or Observational Inadequacy

The codified set of phonetic exponents of filler particles seems to be the vocalic (V) and vocalic–nasal (VN) forms (Lickley 2015) and sometimes also the nasal form (N, *nasal* meaning a nasal consonant). While Smith and Clark (1993, p. 27) add clicks to the set of exponents, this extension is not widely used. Recent research on clicks underpins this point of view, as they are not prototypically used interchangeably with filler particles, but presumably serve different functions in dialogue (for word search and disapproval, cf. Trouvain 2014; for social impropriety, cf. Ogden 2020; for self-repairs, cf. Li 2020). Other potential candidates for a denotational set of filler particles are sniffs, which form "a natural class of delay devices along with objects like *uh(m)* and throat clearings" (Hoey 2020, p. 134). Clicks, sniffs, throat clearings, and "other conduct that borders on the linguistic" are referred to as liminal signs (Dingemanse 2020, p. 191). One of the characteristics of liminal signs together with segmental delay (lengthening) is that they often do not make use of discrete phonetic segments (apart from clicks). Hence, liminal signs are also said to either "preclude speech" (e.g., clicking) or to be "overlappable with speech" (Dingemanse 2020, p. 191). These characteristics are not fulfilled by filler particles, which seem to use at least in parts of their realizations the phonemic inventory of a language and are identifiable as discrete segments without precluding or overlapping speech. Filler particles are therefore not liminal signs.

A common practice when describing the forms of filler particles is to use prototype definitions ('such-as-definitions'), using either phonetic forms "like [əː] and [əːm]" (Trouvain and Werner 2022, p. 62) or orthographically represented forms (in lieu of many: "FP, such as *er, erm, eh* or *ehm*" (Götz 2019, p. 161). Prototype definitions refer to potential linguistic forms of filler particles by listing one or two forms which are prototypical to the class, leaving the other forms unlisted. However, it remains unclear what other potential forms can be subsumed under this definition, as assumptions about the segmental range of the unlisted forms are lacking. For example, in German, the prototypical forms of filler particles consist of vocalic and vocalic–nasal forms, as in many other languages (Lickley 2015). Orthographically represented by *äh(m)*, they nevertheless show substantial vocalic variation in German, subsuming the vowel qualities [œ ɐ ə ɛ] in both vocalic and vocalic–nasal forms (Batliner et al. 1995; Belz 2021; Willkop 1988). In addition, a nasal form exists, orthographically represented as *hm* (for German, cf. Batliner et al. 1995; Belz 2021; de Leeuw 2007; Künzel 1987). Even glottal forms are observed in German, e.g., [ʔəʔ] (Belz

2017). Prototype definitions are therefore problematic for identifying phonetic exponents of filler particles as they run into problems when unconventional forms occur, as further exemplified in Figure 1.
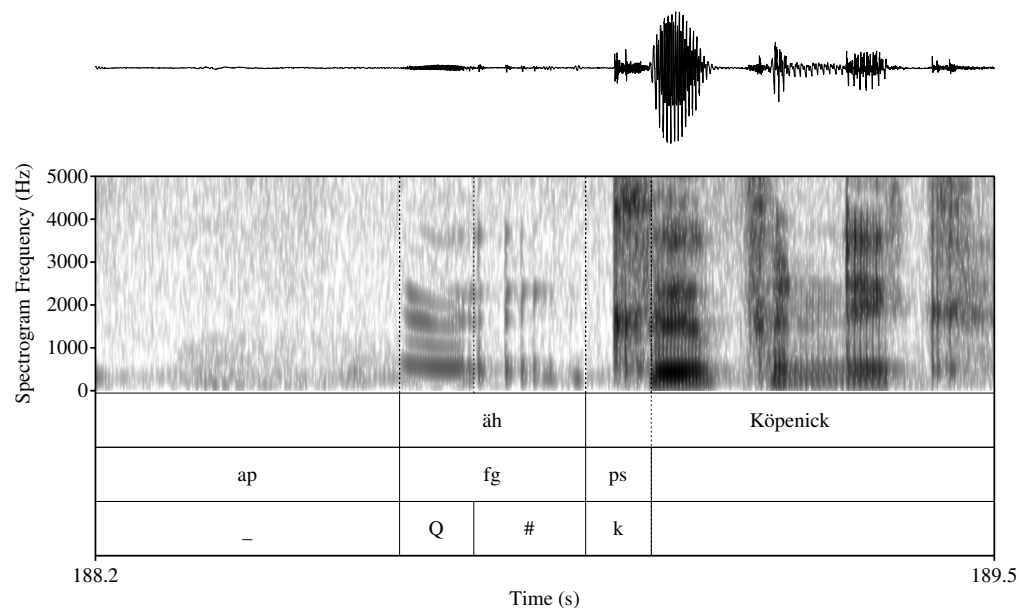

Figure 1. A filler particle (fg) consisting of a squeak (Q) and incomplete vocal fold vibration with low amplitude (#), represented graphematically with *äh*, preceded by a pause (ap) and followed by a segment (ps). *Köpenick* is a German town (source: bh_f_frei_m16).

In Figure 1, the acoustics of *äh*[2] can be described with a squeaky beginning and a glottalized ending, showing incomplete vocal fold vibration with low amplitude. It does not bear any resemblance to the prototype forms V, VN, or N found in German. The evidence for classifying the phonetic exponent as a filler particle is given by the broader preceding context, which is *in Richtung <pause> äh Köpenick* 'in the direction of <pause> äh Köpenick'. The context thus implies a hesitational function. While inserting a filler particle in this context and position seems typical, the acoustic form is unexpected. Nevertheless, in light of the context, position, and function, this specific phonetic exponent needs to be identified as a filler particle. If they had only relied on prototype definitions, however, the annotators might initially not have observed or considered this phonetic exponent to be a filler particle, resulting in a false negative and making any comparison of data across studies and languages less reliable.

As argued above, prototype definitions may result in false negatives when previously unobserved phonetic exponents occur. The reverse option, exhaustive lists, is unmanageable and becomes outdated with the observation of new occurrences. The desideratum identified here is therefore a definition from which potential phonetic exponents of filler particles can be derived together with a term carrying this definition. Although some studies make use of prototype forms as terminological placeholders (cf. for the use of *uh* and *um* Schegloff 2010), this usage can be misleading for the reasons stated above.

2.2. Function—Or Explanatory Inadequacy

Instead of coining a new term for each new concept, existing terms can be given another meaning, a concept called *borrowing* (Mugdan 1990, pp. 54–55). This is what might have happened when one of the first terms for filler particles in modern linguistics, *hesitation-form* (Bloomfield 1933, p. 186), was introduced. Only later did it become clear that "[h]esitation may not always be involved" (Kjellmer 2003, p. 171). In fact, the state of the art is that filler particles are notoriously multi-functional, assuming the (often simultaneous) functions of marking planning at the message-level (Fraundorf and Watson 2013), marking hesitation (Chafe 1980; Corley and Stewart 2008; Goldman-Eisler 1961), marking turns

(Tottie 2015), marking repairs (Belz et al. 2017; Fox Tree and Clark 1997), marking attention (Collard et al. 2008; Schegloff 2010), marking information status (Arnold et al. 2003, 2004), marking complexity of following phrases (Watanabe et al. 2008), marking discourse (Swerts 1998), and marking dialogue structure (Lickley 2001; Nicholson 2007). It follows that with each new function, a new borrowing of terms could happen. From a terminological point of view, the issue is not that filler particles may be used by the speakers of a language with different or overlapping functions in speech, but that terms describing the function of specific phonetic exponents in a specific context are hijacked and then generalized to refer to filler particles as a whole. I will explain this process by using *hesitation/hesitation marker* as an example.

Hesitation can be defined with various degrees of granularity. A general approach describes it as "not know[ing] what to say next (or how to express it)" (Gilquin 2008, p. 120). However, this definition does not help in identifying the phonetic exponents carrying this function. More fine-grained definitions sometimes list specific phonetic exponents, either on a concrete level, as in "a silence of 0.3 to 0.4 s" (Riggenbach 1991, p. 426) (excluding filler particles entirely), or on a more abstract open class structural level, as in "self-repetition, self-correction, or filled pauses" (Baker-Smemoe et al. 2014, p. 714). A recent definition combines functional detail with an open set of phonetic exponents. In this view, hesitation is "anything that temporally extends the delivery of the intended message for whatever reason" (Betz 2020, p. 11). Betz (2020) further defines *hesitation* as "forward-looking disfluency", comprising phenomena such as *lengthening*, *silences*, and *fillers*. The preciseness of the definitions by Baker-Smemoe et al. (2014) and Betz (2020) is improved by the fact that they use the terms *filled pause* or *filler* to refer to potential phonetic exponents, and only in a second step assign the hesitational function to these entities.

Confusion arises when the term *hesitation/hesitation marker* is used for every phonetic exponent found in a corpus. For example, Wieling et al. (2016), in a large-scale study on the variation and change of filler particles, use the term *hesitation marker* to refer to "*er*" and "*erm*". However, it cannot be ruled out that the actual function of a specific *er* used in their data is not hesitational, as the function of filler particles is not systematically investigated in their study. This terminological decision thus obscures terminological preciseness and unambiguousness (cf. Gläser 1995, p. 528) by falsely extending the description of one possible function of a specific phonetic exponent to all studied exponents without examining their actual function.

The examples for *hesitation/hesitation marker* are transferable to other functions of filler particles. Thus, the context-dependent function of filler particles is not a generalizable source for terms to describe the phenomenon precisely and unambiguously as this results in inadequate explanations of phonetic exponents. It therefore seems reasonable to shed light on the level of the linguistic category in the quest for a term that is not prejudiced (leaving room for various functional classifications), a task which is undertaken in the next section.

2.3. Linguistic Category—Or Descriptive Inadequacy—And the Term Filler Particle

Terminological preciseness and unambiguousness, the prerequisites for terms given by (Gläser 1995, p. 528), have so far not been met by terms transferred from form or function, such as *uh(m)* or *hesitation*. I will first discuss the most commonly used terms *filled pause* and *filler* before turning to terms such as *discourse marker* or *interjection* and, finally, making the case for *filler particle*.

The terms *filled pause* and *filler* are often used interchangeably. *Filled pause*, one of the first terms used for describing the phenomenon, is closely connected to the term *silent pause*, which is defined as the absence of linguistic acoustic material (cf. Lickley 2015, but see Belz and Trouvain 2019 for acoustic material within silent pauses). The terminological proximity of *filled pause* and *silent pause* have even misled some studies to make no distinction between the two (e.g., Boomer and Dittmann 1962; Hawkins 1971). *Filled pause*, as opposed to *silent pause*, strongly implies that the identified segment in the speech stream (independent of

its articulatory foundation) is a *pause* which is filled by a vocal event. This results in an oxymoron or misnomer as has often been noted in the literature (e.g., Belz et al. 2017; Ehlich 1986; Eklund 2004; Smith and Clark 1993; Trouvain and Werner 2022). The paradox is due to the fact that a pause is defined as something being absent or interrupted, "a temporary stop or rest, esp in speech or action".[3] However, the acoustic signal of the filler particle gives evidence for articulatory and acoustic activity, so there is de facto no pause in the speech signal. Furthermore, the literature on the phonetic realization of filler particles reports language-, speaker-, and context-specific acoustic characteristics of fundamental frequency (Shriberg 1994; Shriberg and Lickley 1993), vowel formants (Belz 2020; de Boer et al. 2022), and interactions with prosody (Belz 2021; Clark and Fox Tree 2002), corroborating the hypothesis of Gick et al. (2004, p. 231) that filler particles "have targets of their own". Therefore, in a strict articulatory–acoustic phonetic sense, there is no pause (although a listener might nonetheless perceive one). What is meant by the term *pause* in *filled pause*, then, is the absence of a semantically interpretable signal as opposed to pragmatic interpretations. In this sense, the term *filled pause* conveys the notion that filler particles are symptoms, not signals, a discussion that is rekindled time and again.[4] Apart from being imprecise, the term *filled pause* is biased towards a theoretical point of view that does not adequately describe the phenomenon.

The term *filler* was introduced by Clark and Fox Tree (2002).[5] Filler particles are often described as being non-lexical or non-verbal, although the implications of this observation are elusive in the literature. In their attempt to refute the non-lexicality statement, Clark and Fox Tree (2002) investigate silent pauses following filler particles and argue that fillers possess the meaning of "announc[ing] the initiation [...] of what is expected to be a [...] delay in speaking". Grammatically, Clark and Fox Tree (2002) categorize filler particles as interjections, which is a major step in driving filler particles away from the periphery of grammar. However, it has also led to the term *filler* being understood as an interjection *ab initio*, which is problematic in several ways. Although filler particles share some characteristics with interjections, which are also short and uninflected, other criteria are not met. The first argument against the categorization of filler particles as interjections is that interjections "always constitute an intonation unit by themselves" (Ameka 2006, p. 744). While filler particles may constitute intonation units by themselves, this is not always the case (cf. Figure 2 for a standalone filler particle and a filler particle integrated into an intonation phrase).
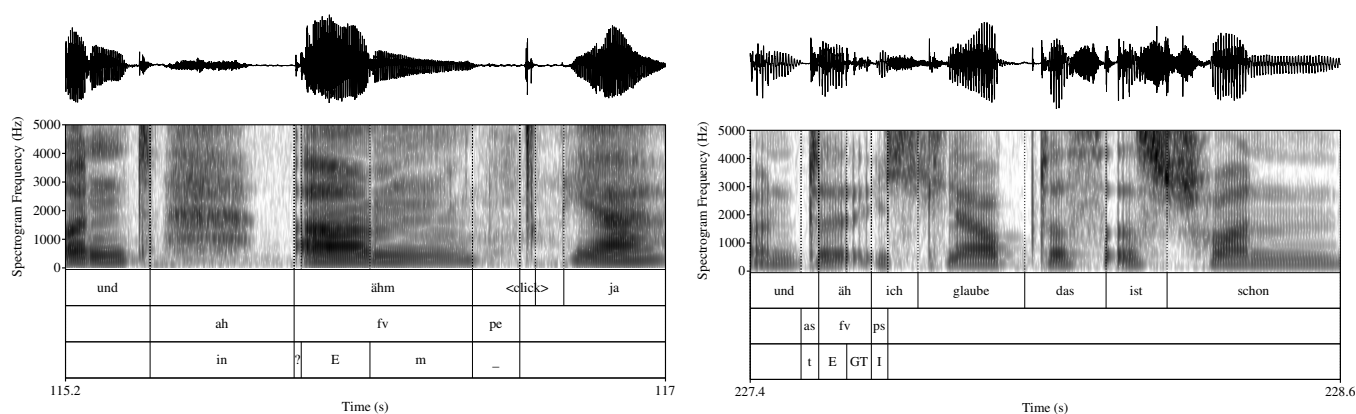

Figure 2. Left panel: Filler particle *ähm* in *und <ah: breathing in> ähm <pe: pause> <click> <pause> ja* 'and <...> ähm <...> yes' constituting a single intonation unit between a breath (ah) and a 'noisy' pause (pe) (source: *bi_f_free_f17*). Right panel: Filler particle *äh* in *und äh ich glaube das ist schon* 'and uh I believe that is already' within an intonation unit, surrounded by phonetic segments (as, ps). The third layer consists of extended SAMPA, cf. Table 2 (source: *bj_f_free_f19*).

The second argument against the categorization of filler particles as interjections is a systematic one. Interjections are categorized by Ameka (1992) into three groups, being

either expressive ("vocal gestures which are symptoms of the speaker's mental state", e.g., *wow!*), conative ("directed at an auditor", e.g., *huh?*, *pst!*), or phatic ("establishment and maintenance of communicative contact", e.g., *mhm*, *uh-huh*, *yeah*) (Ameka 1992, pp. 113–14). For example, the filler particle in the fictional utterance "*um excuse me, could I ask you something?*" might be perceived as an attention marker and thus *establishes* communicative contact, which is why it may be considered to be a phatic interjection (for further examples and an operationalization cf. Section 3, Characterization 5). However, apart from these cases being infrequent, filler particles do not *maintain* communicative contact as required by the description of phatic interjections, as they are not used in a similar position as typical feedback items, e.g., *yeah* or *mhm*. They can even lead to a loss of the communicative contact, acting as turn-closing elements (Tottie 2015). This is corroborated by O'Connell and Kowal (2005), who, in the light of the multi-functionality of filler particles, challenge their classification as interjections. Their argument is a distributional one. Kowal and O'Connell (2004, p. 96) compare the distribution of *äh* and *ähm* in German interviews with that of interjections. They find differences at the end of phrases (final usage of filler particles in 52% of the cases, but of interjections in 7% of the cases), at the start of cited speech (filler particles 2%, interjections 30%), and longer pauses after interjections than after filler particles. They conclude from the different distributions that filler particles are not used in the sense of interjections and they rather argue that they are structuring particles ('Gliederungspartikeln'). The approach of classifying filler particles as particles is therefore not entirely new. To be clear, the analyses as interjection or as particle do not necessarily exclude each other. Rather, I will argue that the phonetic exponent should be seen as a filler particle candidate a priori, after which secondary (pragmatic or functional) analyses can be conducted (see Section 3). This follows from the view that the "functional classification [of an interjection] is based on what is perceived to be the predominant function of the item in question with respect to its semantics" (Ameka 2006, p. 744). As filler particles are semantically empty, their functional classification as interjection might even be a non sequitur.

In conclusion, filler particles do not fulfil any of the classic criteria for interjections, or only do so in specific cases. Using the term *filler* to refer to filler particles might therefore be misleading, as *filler* is defined as an interjection by design. For a possible solution, it is proposed with the definition in the next section that the particle class be considered as a candidate for describing filler particles more adequately.

3. Definition and Operationalization

3.1. The Present Definition

Two terms are defined that can be used to refer to different types of phonetic exponents: filler particle and filler particle candidate. The following definitions come with superscript numbers indicating characterizations explained in detail in Section 3.2.

Definition 1 (Filler particle). [1]*A phonetic exponent* [2]*which is segmentally structured,* [3]*semantically empty,* [4]*syntactically unconstrained,* [5]*and does not show an interjectional function is classified as* [6]*a filler particle.*

Definition 2 (Filler particle candidate). *A phonetic exponent which is segmentally structured, semantically empty, and syntactically unconstrained is classified as a filler particle candidate.*

Definition 1 draws on previous work on filler particles discussed above and refines an earlier suggestion for a definition (Rose 2007).[6] It is an approach to both a theoretically grounded and empirically applicable identification of filler particles in non-pathological speech.[7] It is oriented towards surface fluency, meaning "what the speaker actually produced" and what is "noticeable on close inspection of the speech signal" (Lickley 2015, p. 468).

Definition 2 is part of Definition 1, but ends before a functional classification that rules out interjections. It thus merely defines filler particle candidates (FPC). In corpus studies, this intermediate output can be used to speed up the annotation; the candidates can then be re-evaluated later regarding whether they fulfil the functional prerequisites.

Both definitions aim for a heuristically reproducible identification process for filler particles, so that researchers do not need to rely on prototypical or conventional phonetic forms. In Section 3.2, Definition 1 is explained with respect to its characterizations, together with remarks on filler particle candidates, filler particle form, and filler particle function. Though the definition is later applied to a corpus of German, it aims at ideally being universally applicable (cf. also the discussion in Section 6), which is why cross-linguistic comparisons are included in certain characterizations.

The process of identifying a filler particle is operationalized by the flowchart in Figure 3. Start and end nodes have rounded corners, remarks are symbolized by rectangles, decisions by diamonds and intermediate outputs by rhomboids. To start, a phonetic exponent in speech needs to be noticed (Characterization 1). The first decision is about the segmental structure of the phonetic exponent (Characterization 2). If there are no discrete phonetic segments involved, the phonetic exponent is not a filler particle but something else, e.g., segmental lengthening or a liminal signal. If the phonetic form shows a discrete segment structure, it can then be categorized within a phonetic form (Remark 1). Next, the phonetic exponent is semantically inspected (Characterization 3). If the semantic value in its specific contextual usage is empty, the path continues towards the syntactic constraint (Characterization 4); otherwise, the phonetic exponent is not a filler particle, but something else. If the phonetic exponent is, in principle, unconstrained, i.e., not bound syntactically or taking syntactic scope, it is classified as a filler particle candidate; otherwise, the phonetic exponent is not a filler particle. If the FPC carries interjectional functions in its specific context (Characterization 5), it is not a filler particle. If it clearly has no such function, it is identified as a filler particle (Characterization 6). If no unequivocal decision can be made, the phonetic exponent is allocated to the FPC category. After the identification of a filler particle, its pragmatic functions in speech may be analysed further (Remark 2).

3.2. Characterizations and Remarks

In this section, Definition 1 is broken down into its parts. The text behind each characterization is taken directly from the definition.

Characterization 1 (Signal). → *phonetic exponent*

The phonetic exponent describes the representation of a potential filler particle in the acoustic speech signal. It is a priori different from a speech pause (i.e., the absence of a signal). Phonetic exponents only refer to entities produced by the speaker's larynx and vocal tract, thereby delineating the signal from gestural signs such as finger snaps.

Characterization 2 (Phonematic unit). → *segmentally structured*

Phonematic units are "discrete segmental units [...] that occur in linear sequence" (Brown and Miller 2013). Filler particles consist of phonetic forms qualitatively distinct (i.e., discrete) from the surrounding segmental context in such a way that they are distinguishable from suprasegmental laryngealization[8], vocalic transition, or segmental lengthening. They need to be perceivable both visually and acoustically when inspecting the signal, as shown in the right panel of Figure 2 above. For example, schwa lengthening in *grüne* 'green' in Figure 4 is not classified as a filler particle because there is no discrete phonematic unit distinguishable in the signal.
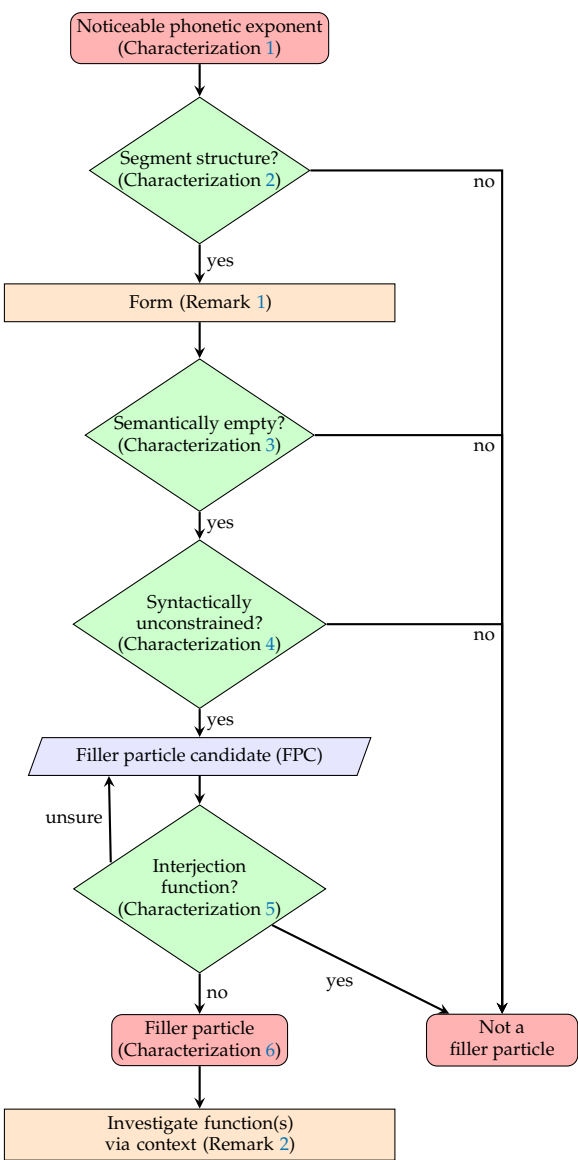
Figure 3. Flowchart for identifying filler particles, with start/end nodes (rounded corners), decision nodes (diamonds), remarks (rectangles), and intermediate outputs (rhomboids).
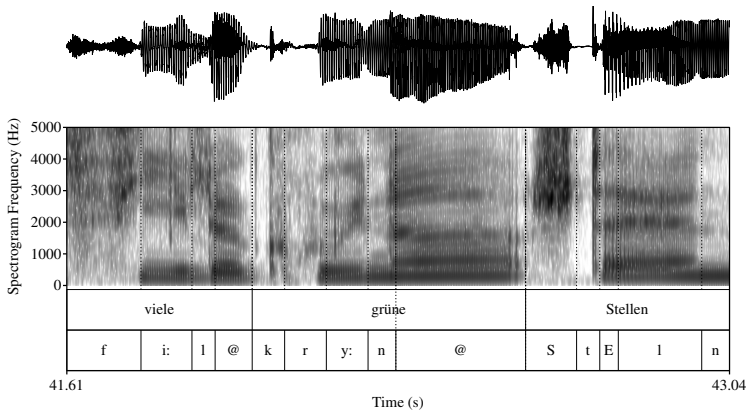

Figure 4. Lengthening of schwa (@) in *grüne* 'green' in the phrase *viele grüne Stellen* 'a lot of green patches', showing a graphematic layer and a segmental layer (source: bj_f_frei_f19).

Further, this characterization also distinguishes filler particles from liminal signs, including non-verbal vocalizations (cf. Section 2.1). Liminal signs or non-verbal vocalizations are defined by Dingemanse (2020) as "conduct that borders on the linguistic". Liminal signs do not use the phonemic inventory of a language (except for clicks).

Remark 1 (Form). *Though typically one to three prototypical filler particles per language are (phonologically) observed (e.g., Lickley 2015), there can still be substantial (phonetic) variation, which is why there is no a priori list of expected phonetic exponents given here. For example, for German, the forms [ə ɐ ɛ œ ø] and [əm ɐm ɛm œm øm] (Batliner et al. 1995; Belz 2020), which vary in vowel quality, as well as glottal forms (Belz 2017) are observed. In addition, up to this point, clicks are still potential candidates for filler particles.*

Characterization 3 (Semantics). → *semantically empty*

This characteristic is adopted from Rose (2007) and distinguishes filler particles from other elements such as lexical items. However, the notion of lexicality is not integrated in the definition of filler particles (by using, e.g., a phrase such as "non-lexical entity"). The reason is the trouble with the term *lexical*, which is used ambiguously in linguistics, referring both to words "denoting things, beings, events, abstract ideas and so on" and "the vocabulary of a language, as part of its description" (Brown and Miller 2013). While filler particles may be described as non-lexical in the sense that they do not denote things, beings, events, or abstract ideas, the reading of *non-lexical* in the sense of not being in the vocabulary of a language is misleading, as dictionaries have started to list filler particles (at least orthographically)[9]. However, if one dismissed the rather weak argument that being represented in a dictionary entry is proof of lexicality, the Spanish filler particle *esto* (as listed in the dictionary) would still stand out from prototypical vocalic or vocalic–nasal phonetic forms. While it might be argued that *esto* is a demonstrative pronoun used in a pragmatic filler function and can therefore be considered to be lexical, Graham (2018) counters this argument, reasoning that the usage of *este* (*sic*, as observed phonetically, not *esto* as listed in the dictionary) is different from its demonstrative usage, as the phonetic form of *este* shows "customary elongation of the final syllable" and there is no observable agreement with other constituents (Graham 2018, p. 2).

Characterization 4 (Syntax). → *syntactically unconstrained*

Filler particles are constituents that are syntactically unconstrained, meaning that no restrictions on placement apply. This can be validated by using linguistic constituency tests, e.g., omission, movement, or general substitution. If the grammaticality of the sentence or utterance is not affected, no restrictions apply. Nevertheless, filler particles can show distributional characteristics. In English and French, they tend to occur frequently at the beginning of phrases and are rarely found within collocations or chunks (Crible et al. 2017; Schneider 2014). An example of a syntactically constrained but semantically empty element (at least it is described as such) is the German adverb *gar*, which functions as an intensifier.[10] It is, however, syntactically constrained (as validated by the ungrammatical outcome of moving it to another position) and is thus not identified as filler particle.

Characterization 5 (Interjection). → *does not show an interjectional function*

Conative, expressive, and phatic functions are integral to interjections (cf. Section 2.3 and Ameka 2006). If filler particle candidates fulfil one of these perceived functions, they are then classified as interjections. For example, the FPC *eh* in Spanish is used as an interjection, an interrogative marker, and a filler particle (Roggia 2012). When used as interrogative marker, *eh* shows a rising intonation at the end of an utterance (Roggia 2012, p. 1787), therefore exhibiting a conative function, and is thus not being used as filler particle. In

the same way, the phonological string *gar* in German functions as either an adjective or an adverb depending on the context. It follows that [ʔɛː] may function as both a filler particle and an interjection depending on the context and its phonetic features. Thus, anecdotally, *äh* in the ironic–sarcastic feedback construction *äh genau* 'uh right', which can serve as a disbelieving reaction to a farcical claim of an interlocutor, might be perceived as an interjection, as it has an expressive quality. While an expressive function can be determined via marked phonetic characteristics such as an unusual pitch contour (cf. Section 5.1.3), it is open to debate whether these instances are perceived as interjections. In unsure cases, the most comprehensible option is to classify them as FPC.

Clicks are notoriously difficult elements of speech, as they may occur as by-products of swallowing, saliva production, or a higher intra-oral pressure due to constriction in the nasal cavity, for example because of a common cold. Clicks are often used in an expressive way in languages where they occur paralinguistically, i.e., outside of the phonemic inventory of the language (for German, cf. Trouvain 2015). Even in the context of other filler particles, clicks may be interpreted as having an expressive quality, e.g., demonstrating the annoyance of not being able to remember a word fast enough (cf. Figure 5), which is conceptually comparable to clicks in self-repairs (Li 2020). There, the FPC click is embedded in the context of the speaker trying to remember a certain place, producing *ah ich ähm hab die in Erinnerung <pause> irgendwo <pause> <click> <pause> ähm <pause> in der Nähe vom* 'uh I um remember <pause> somewhere <pause> <click> <pause> um <pause> near the'. Although clicks fulfil the prerequisites for FPCs and are often used in turn-initiations, it is unclear whether an expressive quality is automatically attributed to their appearance (which would align with the function of social impropriety hypothesized by Ogden 2020), which might be a consequence of their non-phonemic status in German. Clicks therefore warrant further research before they can be classified unequivocally as filler particles—in this study they are therefore regarded as FPCs and not investigated further.
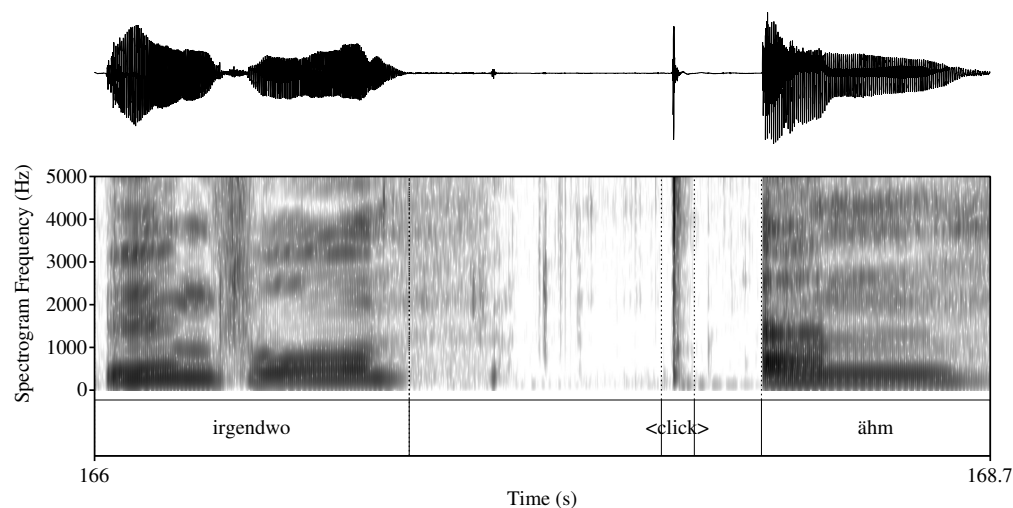

Figure 5. A click before the vocalic filler particle *ähm* (fv) in *irgendwo <click> ähm* 'somewhere <click> um' (source: bh_f_frei_m15).

Other cases of doubt are truncations. For example, a singleton [f] can either be analysed as a remnant of a truncated word or as a filler particle. If the utterance context includes a repair and a word starting with [f] is uttered within the right syntactic structure, it could very well be a truncation and would be excluded by Characterization 3. Otherwise, without traces verifying a truncation, it needs to be analysed as a filler particle.[11] It is open to further discussion whether target hypotheses should be included in the analysis of filler particles. At this point, I propose that phonetic exponents which cannot be ruled out as filler particles should be analysed as filler particles.

Characterization 6 (Part of speech). → _A filler particle_

　　Terminologically, _filler particle_ consists of _filler_ and _particle_. The first element is borrowed from the history of filler particle research, as it is recognizable, but does not carry a prominent pragmatic function biasing the term. The second element _particle_ specifies that the phonetic exponent is part of the particle class. It is therefore by definition not a pause, but a word, from which follows that filler particles are pause-external entities.

　　Filler particles are classified within the part-of-speech class of particles. Particles are understood as "small, uninflected words that are only loosely integrated into the sentence structure, if at all" (Fischer 2006, p. 4) and which do not fit a standard classification of parts of speech (Brown and Miller 2013; Crystal 2008). The subsuming of fillers under the particle class was inspired by Keseling (1989), who is, to the best of my knowledge, one of the first authors to use _particle_ to describe filler particles. Other studies that consider fillers as particles include Kowal and O'Connell (2004); Willkop (1988) (structuring particle), Klug (2013) (hesitation particle), Belz (2021) (filler particle), and Trouvain and Werner (2022) (phonetic particle) for German, Roggia (2012) (discourse particle) for Dominican Spanish, and Kirjavainen and Nikolaev (2022) (planning particle) for Finnish. The term _marker_ is used differently from particle, as it refers to the specific function of a phonetic exponent in context. For example, the term _discourse marker_ refers to words or multi-word expressions that show a specific function in discourse, but are in other ways understood as adverbs or particles. In comparison to discourse markers, which are multi-functional, syntactically optional, and have a fixed form and a variable scope (Crible 2017, p. 106), filler particles differ in that they do not take scope.

Remark 2 (Function). _The definition applied here distinguishes between a primary functional difference of the usage in context which reflects back to word class, and secondary functions in the context of a multi-functional paradigm of filler particles. If a filler particle candidate can be shown to carry expressive, conative, or phatic features (Ameka 2006), it may be classified as an interjection and, otherwise, as a filler particle. The broadly discussed functions of filler particles (e.g., turn-taking, turn-yielding, hesitation marking, marking of information structure) may then be attributed in a second step._

3.3. Hypotheses

　　The definition was constructed with a focus on finding all true positive instances of filler particles in the data. It thus should have two consequences:

Hypothesis 1. _The definition should yield higher frequencies in the data than previously found for a specific corpus that had been annotated without the explicit definition presented here._

Hypothesis 2. _The definition should find forms that have to date not been represented orthographically, i.e., non-prototypical filler particles._

4. Materials and Methods

4.1. Method

　　To empirically underpin the definition argued for here, a quantitative–qualitative approach is taken. Therefore, a corpus-based study has been conducted to apply the definition. Some exemplary cases of prototypical and non-prototypical filler particles are presented and described. To test Hypothesis 1, a subcorpus of BeDiaCo (s. Section 4.2) has been re-annotated for filler particles in German and the outcomes compared to each other. To test Hypothesis 2, the segments of filler particles in the same subcorpus have been analysed.

4.2. BeDiaCo Corpus

The Berlin Dialogue Corpus (BeDiaCo, Belz et al. 2021) was used. It is available for linguistic research at the media repository of Humboldt-Universität zu Berlin[12] and is documented extensively (Belz et al. 2021). For this study, 32 native speakers of German in 16 dialogues were used.[13] The subcorpus BeDiaCo-main with 16 speakers (8 dialogues) was used to test Hypothesis 1. Hypothesis 2 was tested for 32 speakers, including the copresent condition of the subcorpus BeDiaCo-videocall to keep the setting comparable (in both subcorpora, speakers were recorded face-to-face in a sound-attenuated booth). The subcorpora differ in some small features, as BeDiaCo-videocall was recorded during the COVID-19 pandemic and for other research hypotheses as well. Therefore, in BeDiaCo-main, speakers who did not know each other were told to speak freely for ca. 15 min about a topic of their choice, suggesting food as a conversation starter, whereas in BeDiaCo-videocall, speakers who knew each other were told to speak freely for ca. 10 min about the city of Berlin or about their dream vacation. Table 1 presents the corpus summary.

Table 1. Corpus summary of BeDiaCo.

	Subcorpus Main	Subcorpus Videocall	Total
Number of diplomatic tokens	25,269	16,812	42,081
Time frame of single dialogue (min)	15	10	
Duration of articulation (min)	115	73.8	188.8
Participants	16 (aged 18–31)	16 (aged 19–32)	32
Gender	10 m, 6 f	8 m, 8 f	18 m, 14 f
Language		Native speakers of German	
Topic	Free	Free	
Conversation starter	Food	Berlin or dream vacation	
Relationship	strangers	siblings, flatmates, partners	

4.3. Annotation, Query, and Statistics

BeDiaCo uses a flexible multi-layer architecture encompassing both the acoustic signal and annotation layers. Filler particles are annotated graphematically on the diplomatic transcription tier `dipl` (meaning that the transcription does not need to follow the orthographic norm), categorized for form on the filler particle tier `fp`, and segmentally annotated on the segmental tier `segm`. For this study, clicks were omitted. Table 2 summarizes the annotation tiers and their possible values.

The annotation was conducted manually in Praat (Boersma and Weenink 2022). Every sound file of every speaker was played from the start to the end. All phonetic exponents as identified by the definition were coded as filler particle or filler particle candidate on the `FP` tier. Boundaries were annotated in Praat where the sound is only just or just no longer identifiable as a sound of a particular class. Boundaries were set at the zero crossing of rising slopes in the oscillogram. The data were then converted into an EMU data base (Winkelmann et al. 2017) with the R package *emuR* (Winkelmann et al. 2020).

Corpus queries and statistics were conducted in R (R Core Team 2022). To calculate filler particles per hundred words (phw) and per minute, all tokens on the diplomatic tier `dipl` were queried, excluding pauses, laughter, clicks, and non-identifiable passages. Filler particles per minute thus refers to the articulation time of all tokens per speaker.

Table 2. Annotation values for the three annotation layers *dipl, fp, segm* in BeDiaCo.

Tier	Value	Description
`dipl`	*	Open class, e.g., *äh*, *ähm*, *hm*.
`FP`		A filler particle or filler particle candidate as defined in Definition 1 and 2.
	FP	A filler particle.
	FPC	A filler particle candidate.
`fp`		A filler particle as defined in Definition 1.
	fv	A vocalized filler particle.
	fg	A glottal filler particle without a perceivable vocalic structure.
	fx	Not categorizable, e.g., pseudonymized signal at this interval or unclear whether it consists of a truncation or a lexical word unknown to the annotator.
`segm`	*	Open class segmental annotation using SAMPA, e.g., [?Em].
	?	Comprises one to three glottal stops closer than 50ms from each other.
	G	Comprises more than three glottal stops closer than 50ms from each other.
	GP	Sequence of more than two glottal stops farther away than 50ms from each other.
	GT	Glottalized transition between two vowels. The vowel boundary in *fv* is chosen as the start or end of modal phonation.
	GG	Glottalized sequence uttered with the mouth closed.
	Q	Compressed voice with extremely high fundamental frequency.
	E	Vowels in *fv* are always annotated as *E*, the interval ending after the last complete vowel period.
	#	Incomplete vocal fold vibration with low amplitude, may occur before or after vowels.
	x	Indistinguishable segments.

5. Results of the Corpus Study

5.1. Examples of Identifying Filler Particles

5.1.1. A Prototypical Filler Particle

Figure 6 shows a vocalic filler particle. Its identification procedure was as follows. First, the input was recognized as segments not belonging to the left (*und*) or right (*genau*) context (Characterization 2), leading to a vocalic phonetic exponent that is also semantically empty (Characterization 3) and syntactically unconstrained (Characterization 4). The resulting filler particle candidate is not used in a conative, expressive, or phatic way in this context (Characterization 5) and was thus identified as a filler particle.
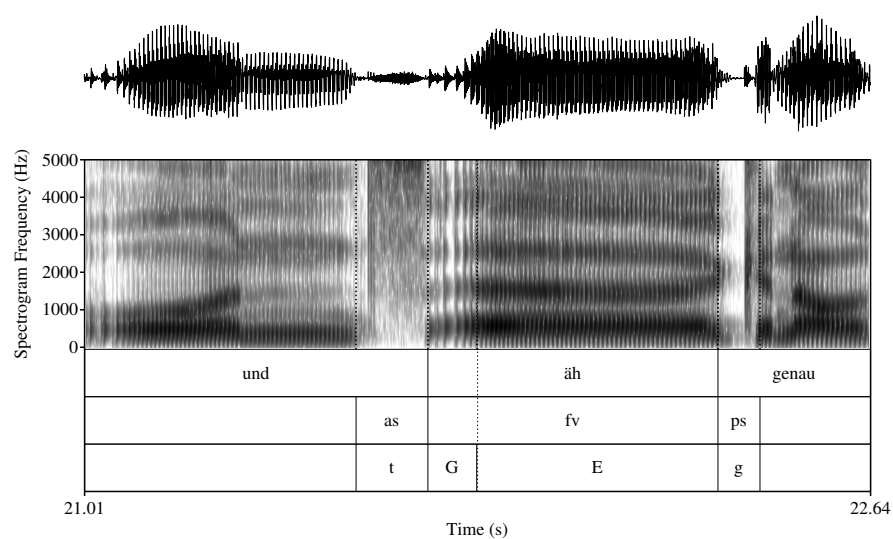
5.1.2. Non-Prototypical Filler Particles

Figure 7 shows a glottal filler particle. Following the flowchart in Figure 3, the input was recognized as segments not belonging to the left (*ich*) or right (*also*) context (Characterization 2). The four glottal stops after [ç] were perceived differently both in phonation and vowel quality from the [a] vowel of *also*. Although the formant structure for F1 and F2 can be conjectured, it was classified as a glottal form due to a lack of higher formant structures. The phonetic exponent is both semantically empty (Characterization 3) and syntactically unconstrained (Characterization 4). The resulting filler particle candidate is not used in a conative, expressive, or phatic way in this context (Characterization 5) and was thus identified as filler particle.

Figure 6. Example of a filler particle *äh* (602 ms) in *und äh genau* 'and uh right', with a left segmental context (as) before the vocalic filler particle (fv), followed by a right segmental context (ps). G denotes a glottalized sequence, E the vowel (source: n_frei_m15).
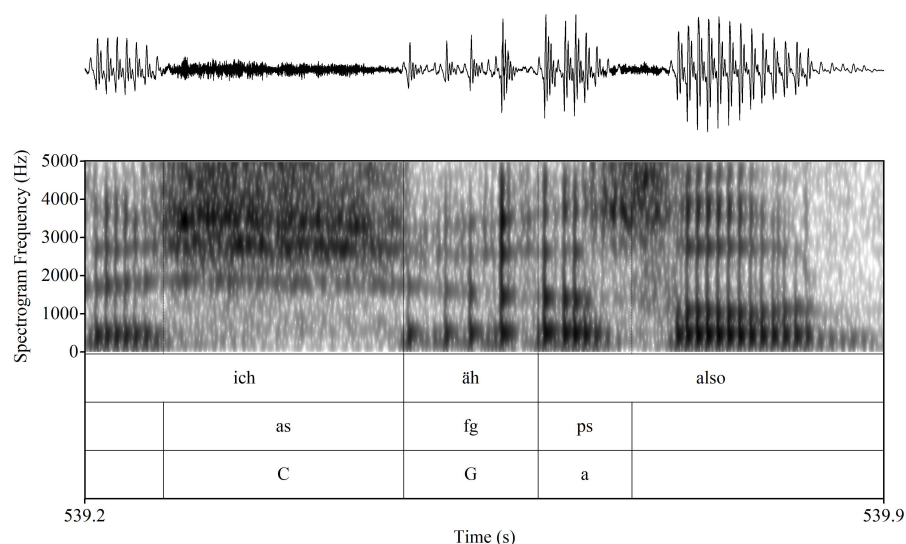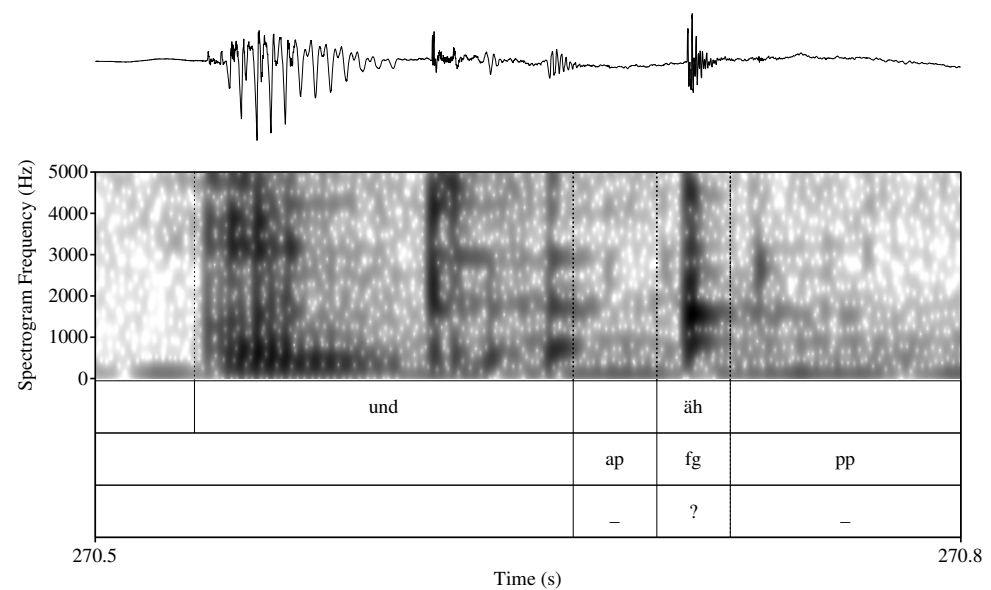

Figure 7. Example of glottal filler particle *fg* (112ms) being categorically different from the following [a] in the postsegment ps (source: bf_f_frei_m12).

Figure 8 shows a glottal filler particle consisting of a singleton glottal stop [ʔ]. Following the flowchart in Figure 3, the input was recognized as segments not belonging to the left (*und*) or right context, which is a pause (Characterization 2). Glottal stops are known to co-occur epiphenomenally after the release of a consonant closure, as is the case just before *ap* in Figure 8. For the singleton glottal stop in *fg*, however, the following pause and the high amplitude of the burst render it less plausible to assume that this case is merely a passive corollary of the previous alveolar stop. The phonetic exponent is both semantically empty (Characterization 3) and syntactically unconstrained (Characterization 4). The resulting filler particle candidate is not used in a conative, expressive, or phatic way in this context (Characterization 5) and was thus identified as a filler particle.

Figure 8. Example of glottal filler particle *fg* consisting of a singleton glottal stop [ʔ] (23 ms) between two pauses ap and pp and following the conjunction *und* (source: bb_f_frei_f3).

Figure 9 shows the vocal filler particle [jɛːvə]. Following the flowchart in Figure 3, the input was recognized as segments not belonging to the left (*geschmacklich*) or right (*besser*) context (Characterization 2). The form consists of two vowels. The phonetic exponent is both semantically empty (Characterization 3) and syntactically unconstrained (Characterization 4), as validated by constituency tests. For example, omitting *jähve* does not render the sentence ungrammatical. Similarly, it could be moved to the left (forming *jähve geschmacklich*) without rendering the sentence ungrammatical or changing its semantics. The resulting filler particle candidate is not used in a conative, expressive, or phatic way in this context (Characterization 5). It is used in the production *wenn tatsächlich das das Geheimnis ist dass das so äh so <pause> geschmacklich jähve besser wird wenn das ein mehrstufiger Sauerteig <pause> is sollten wir das mal probieren* 'if in fact that is the secret that makes it such uh such <pause> better in taste [jɛvə] if that is a multistage <pause> sourdough we should try it some time' and was thus identified as a filler particle.
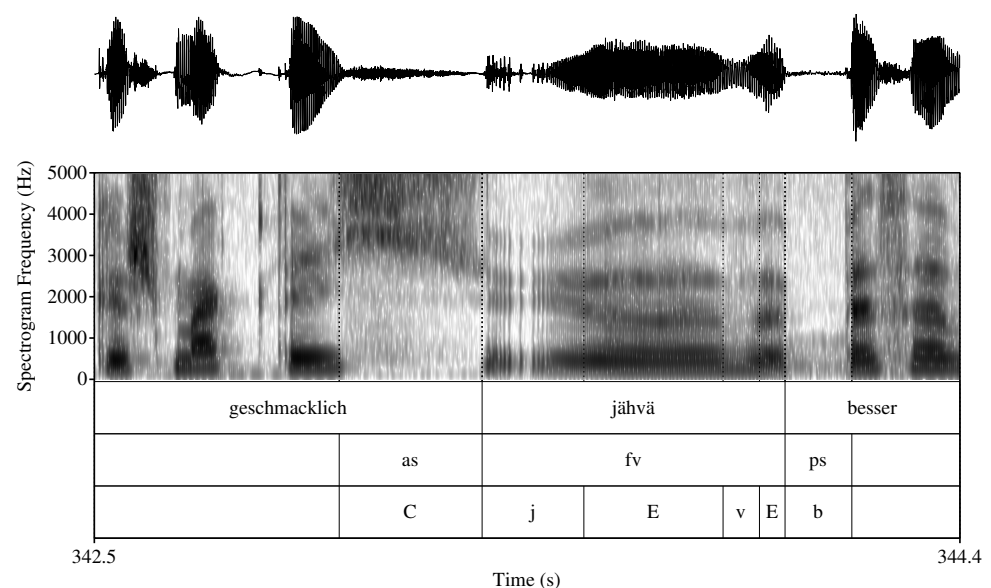

Figure 9. Example of a vocalic (fv) filler particle [jɛːvə] (674 ms) between two segments (bf_f_frei_f11).

5.1.3. Not a Filler Particle, Unsure If an Interjection

Figure 10 shows the phonetic exponent *mh* [m:], a form which sometimes occurs as a filler particle in German (cf. Section 5.2). In this context, however, it was not identified as a filler particle. Following Definition 2, it shows a discrete segment (Characterization 2), is semantically empty (Characterization 3), and is syntactically unconstrained (Characterization 4). It was ruled out as a filler particle, though, as it is used expressively in the context at hand (Characterization 5), as indicated by its realization with an extremely high pitch accent. Paraphrasing its occurrence in the context *und ich wollt dann noch meinen Führerschein machen mh <pause> schaffen wir* ('and then I wanted to get my driver's license mh <pause> we can do that'), it may be understood as saying "I don't know if this might happen in time but we will see". Its expressive function thus suggests that it might be used as an interjection and not as a filler particle. However, as this is not entirely clear, this exponent is classified as FPC.
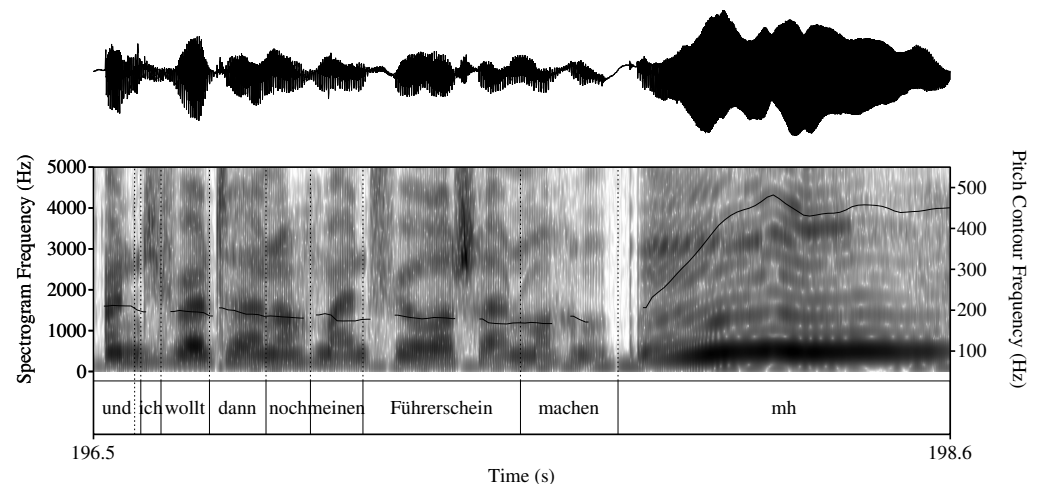

Figure 10. Example of *mh* (687ms) used as an interjection in *und ich wollt dann noch meinen Führerschein machen mh* 'and then I wanted to get my driver's license mh' (source ba_f_frei_f2).

5.1.4. A Filler Particle Candidate

Similar but different to the *mh* analysed above is the instance of *hm* [m:] as shown in Figure 11. After the statement of speaker m.m4 *aber da muss man halt immer sofort Zeit haben wenn da die Email kommt so* ('but you always have to have time immediately when the email arrives'), speaker m.m5 replies *ach so <pause> hm ja okay* ('I see <pause> hm yeah okay'). Following Definition 2, the phonetic exponent of *hm* shows a discrete segment (Characterization 2), is semantically empty (Characterization 3), and is syntactically unconstrained (Characterization 4). However, it cannot be ruled out as a filler particle, as it is ambiguous concerning its communicative intent. Due to the slight variation in the fundamental frequency, it is unclear whether an evaluation of the previous statement by speaker m.m4 is given here. As no clear decision can be found following Characterization 5, this instance is categorized as FPC.
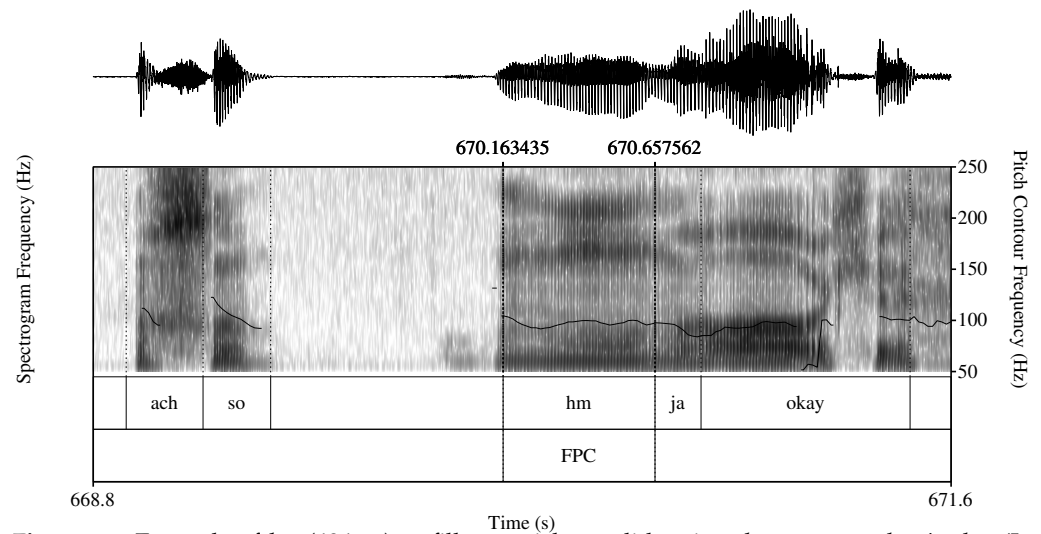
Figure 11. Example of *hm* (494ms) as filler particle candidate in *ach so <pause> hm ja okay* 'I see <pause> hm yeah okay' (source g_frei_m5).

5.2. Frequencies and Forms

After excluding four instances of *fx* on the `fp` tier, 36 participants in 6.8 h of free speech produced a total of 735 filler particles, consisting of 35 glottal and 700 vocalic forms. On the FP tier, a total of 746 entities were labelled. Of these, 735 entities were identified as filler particles and 11 as filler particle candidates. This means that without the FPC category the Type 1 error rate (false positives) would have amounted to 1.47%.

The frequency of filler particles in this corpus is rather low (cf. Table 3), but not unusual for free conversation with another person without solving a specific task. Figure 12a) shows the values for vocalic and glottal filler particles (FP) per hundred words per speaker as well as b) a bootstrapped mean with a 95% confidence interval. The distributions of both vocalic and glottal filler particles per speaker are not normally distributed. This is due to speaker m.f13 for vocalic filler particles (producing more than 11 vocalic FP/min) and due to speaker m.f1 for glottal filler particles (producing more than 1 glottal FP/min). Although these data points present outliers in a technical sense, they are not excluded here as they are not attributable to false positives.

Table 3. Median, mean, and standard deviation (sd) in BeDiaCo v3.

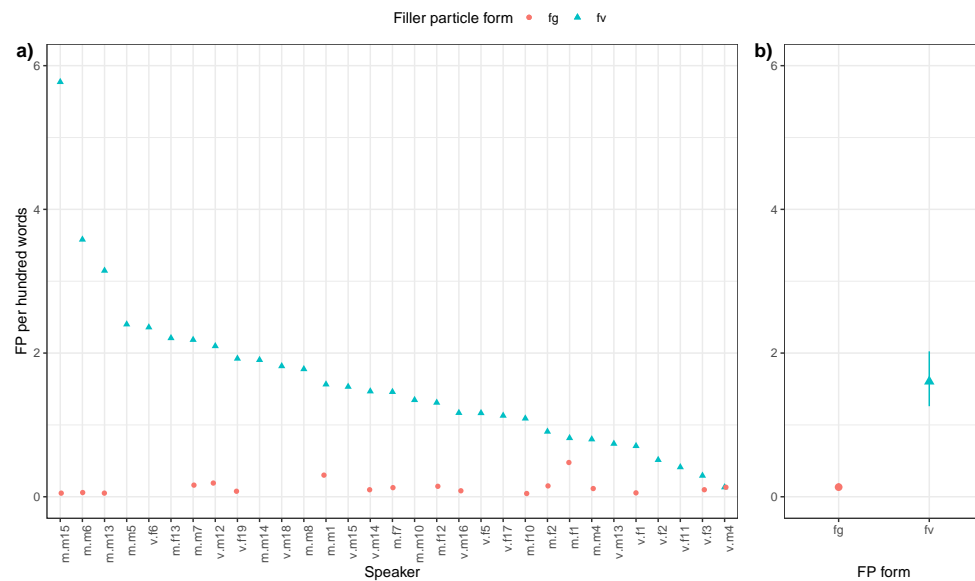| Labels	Median	Mean	sd	Unit
fg	0.25	0.29	0.22	per minute
fv	3.18	3.50	2.25	per minute
fg	0.11	0.13	0.11	per hundred words
fv	1.46	1.60	1.11	per hundred words

Figure 12. (**a**) Filler particle forms (*fg* = glottal, *fv* = vocalic) per hundred words per speaker and (**b**) bootstrapped confidence intervals (95%). The first character (m/v) indicates the main or videocall subcorpus.

Comparing the results of this analysis of BeDiaCo with the same 16 speakers as in an earlier version of the corpus, BeDiaCo v1, (486 vocalic, 21 glottal FPs), the re-annotation with the refined definition of filler particles did in fact reveal more instances in v3 (517 vocalic, 27 glottal FPs), although the difference is not significant, as validated by a χ^2 test.

Turning to the segments of the identified vocalic FP forms, Figure 13 depicts the relationship between segment sequences and orthographic representations in the diplomatic tier for the most common ($n > 2$) segment sequences in the corpus. These instances are represented orthographically by *äh*, *ähm*, *f*, and *hm*. Some 74 cases of category *fv* only occur twice or less in the corpus (cf. Table 4 for all 13 cases with two instances), while 61 *fv* cases occur only once in the corpus.
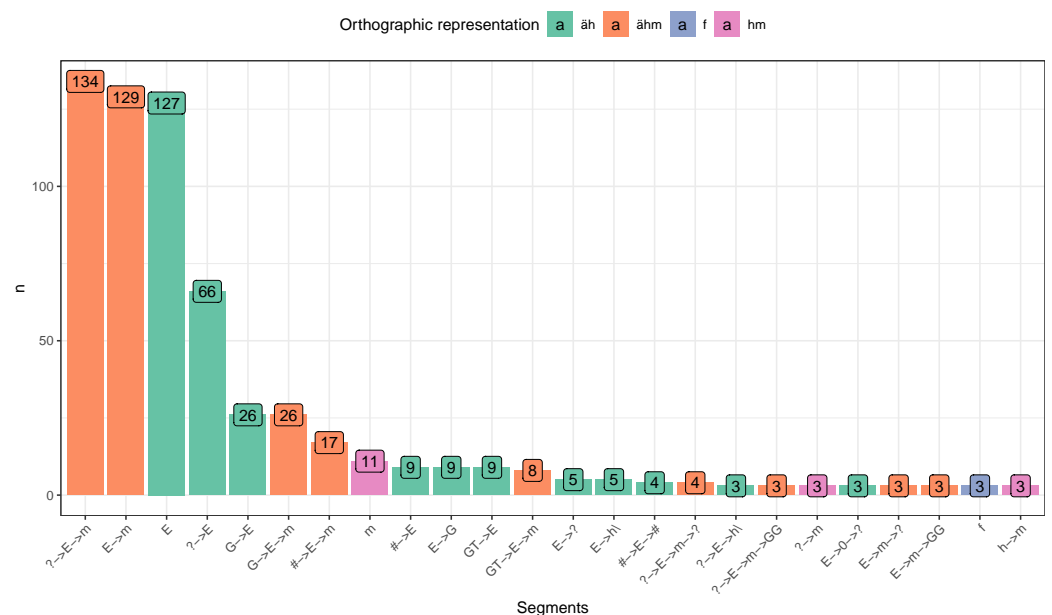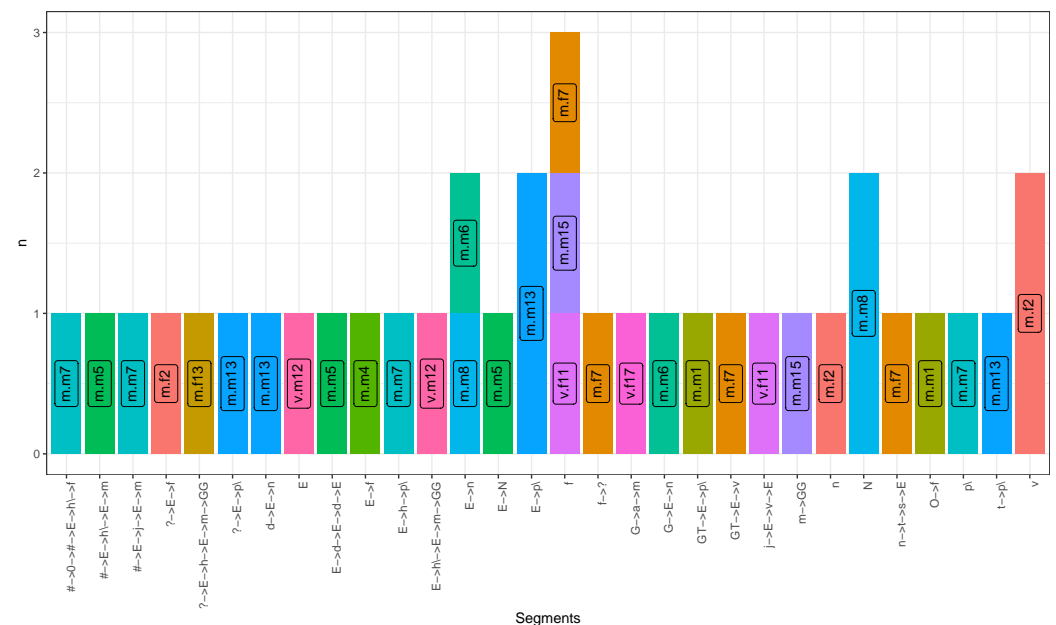

Figure 13. Segments with $n > 2$ instances of the vocalic (*fv*) form per sequence and orthographic representation.

Table 4. Segment sequences for vocalic filler particles with *n* = 2.

Labels	dipl
#→E→G	äh
?→E→?	äh
?→E→GT	äh
E→GP	äh
E→GT	äh
E→h	äh
G→E→?	äh
?→#→E→m	ähm
E→m→G	ähm
G→E→m→GG	ähm
GG	hm
N	n
v	w

There are sequences that deviate substantially from the prototypical forms in German, which would falsely be represented orthographically by *äh*, *ähm*, and *hm*, for example, j→E→v→E (*jähvä*, cf. also Figure 9), E→d→E→d→E (*ähdähdäh*), E→f (*ähf*), and n→t→s→E (*ntsäh*). Figure 14 shows these 36 non-prototypical instances, which represent 5.1% of all vocalic forms. There are speaker-specific cases (e.g., [v] is only used by speaker m.f2), but also identical sequences that are used by different speakers (e.g., sequences with a fricative or approximant at the end are used by m.m1, m.m4, m.m7, m.m13, m.m15, m.f2, m.f7, and v.f11), hinting at a substantial and systematic occurrence in the speech community.

Figure 14. Non-prototypical segments of the vocalic (*fv*) form per sequence and speaker given in SAMPA (p\ = [ɸ], h\ = [ɦ]; for other symbols, cf. Table 2).

Glottal filler particles are much rarer than vocalic filler particles, with only 35 instances in the whole corpus. A peculiar finding is that there are 27 instances in BeDiaCo-main (with 16 speakers), but only 8 instances in BeDiaCo-videocall (with 20 speakers). Segment-wise, Table 5 shows that in half of all glottal filler particles the segment *G* occurs, which indicates that more than three glottal stops closer than 50 ms from each other are produced. The second most frequent segment is a singleton glottal stop (for a visual example, cf. Figure 8).

Other segments occur only once, such as a squeak (Q) followed by incomplete vocal fold vibration with low amplitude (#) (for a visual example, cf. Figure 1).

Table 5. Segment sequences for glottal filler particles.

Labels	n
G	17
?	7
GP	4
?->?->?	1
d->E	1
E	1
G->?	1
G->v	1
n	1
Q->#	1

Almost none of the segmental sequences given in Table 5 can be represented orthographically. Exemptions might be "E", "d→E," and "n". These annotation tags were chosen as the segments are strongly glottalized but still carry perceivable (at least to the annotator) vocalic or nasal features.

Assuming that non-prototypical forms in German are different from the phonetic exponents of the phonologically occurring /ɛ ɛm m/ and consist of glottal filler particles as well as multi-syllabic vocalic forms and mono-syllabic vocalic forms ending in consonants (thus excluding [?]), the proportion of non-prototypical forms (n = 70) amounts to 9.5% of all filler particles (n = 735). This difference is significant (Exact binomial test, $p < 0.001$), thus also serving as an indirect validation measure for the annotation guidelines.

6. Discussion

6.1. Summary

This study addressed the terminological conundrum regarding the term *filled pause* (and others) and instead proposed the term *filler particle* together with a formal part-of-speech definition which can be applied to identify filler particles in the acoustic signal. Hypothesis 1 (*the definition should yield higher frequencies in the data than previously found for a specific corpus that had been annotated without the explicit definition presented here*) was not confirmed. Although the re-annotation with the refined definition yielded higher frequencies, the difference was not significant. One reason for this could be that the definition was already implicitly applied in Version 1 of the corpus, though it had not yet been formalized in the way proposed here. Since no direct comparison could be made between the guidelines presented here and conventional guidelines, the frequencies of non-prototypical phonetic exponents were compared with prototypical ones as found using conventional guidelines. When evaluating this proxy variable, a significantly higher number of filler particles was found. Hypothesis 2 (*the definition should find forms that have to date not been represented orthographically, i.e., non-prototypical filler particles*) was confirmed and replicated for both previously analysed data in the main subcorpus and new data in the videocall subcorpus for free dialogues.

6.2. Linguistic Category

The question of the linguistic status (are they similar to words?) of filler particles has been around for a long time in the literature (Clark and Fox Tree 2002; Keseling 1989), together with the question of whether they are rather symptoms or signals (Finlayson and Corley 2012; Walker et al. 2014), or both (Reitbrecht 2017), or both depending on their context (Kosmala and Crible 2022), being used either as fluent or disfluent constructions. At the same time, Kosmala and Crible (2022) argue against a word status of filled pauses. While the approach taken here does not claim that there is an unequivocal solution to the

symptom–signal problem, it does, however, argue that phonetic exponents falling under the proposed definition that consist of discrete acoustic material noticeable in the speech signal can be described and classified linguistically as particles. While in the view of Kosmala and Crible (2022) "[filled pauses] are [...] still widely different from 'words' because of their extreme mobility and the difficulty to pin down their meaning", the point of view advocated for here objects to this conclusion. Though highly variable and not syntactically bound in principle, filler particles do, just as other particles, show certain distributional characteristics, e.g., with respect to surrounding words or pauses (Clark and Fox Tree 2002; Crible et al. 2017; de Leeuw 2007; Jessen 2012; Rose 2015), syntactic phrases (Bada and Genç 2008; Shriberg 1994), intonation phrases (Peters 2005; Swerts 1998), and dialogue structure (Lickley 2001; Nicholson et al. 2010; Rendle-Short 2004). Even further support comes from a recent study which shows that the position of filler particles affects sentence recall, with the authors concluding that "filled pauses are similar to grammatical items such as suffixes, clitics or prepositions in that their location within a sentence is relatively rigid" (Kirjavainen et al. 2022). The particle class as a functional grammatical class fits the distributional observations in the literature quite well, although I have reservations regarding the categorization of filler particles as suffixes, since suffixes, by definition, do not occur freely.

The similarity to clitics as mentioned by Kirjavainen et al. (2022) might be a fruitful approach for the interpretation of glottal filler particles, as clitics refer to the reduced form of something larger. In the following, I will argue that filler particles can occur on a continuum from phonetic segments to full-fledged multi-syllabic words. Looking at the segmental sequences of filler particles in more detail, both 'small' forms consisting of one or two glottal stops and 'large' forms consisting of multi-syllabic forms (e.g., [jɛvə]) are observed in the investigated corpus. Two categories were used to categorize them, namely glottal and vocalic FPs. Glottal FPs consist of segments produced mainly with the help of the glottis that do not show clearly defined vowel characteristics (e.g., a high amplitude and formants). However, this categorization might also fail when a strongly glottalized vowel is observed (due to the vowel features, such an instance would be a vocalic FP). Why are these two different forms produced by speakers, and how can glottal forms be explained? One hypothesis built from this evidence is that filler particles may draw from a continuum of forms from one or two glottal stops to concatenated multi-segmental forms. This continuum keeps filler particles flexible for insertion into the speech stream for a smaller or larger amount of time and thus can be exploited by the speaker depending on the conversational context.

The continuum hypothesis could only be generated thanks to the phonetic definition of filler particles, which is free from a written language bias (Linell 1982) and does not exclude any phonetic exponents a priori. The continuum hypothesis enables a novel perspective on filler particles. A visualisation with examples from German is given in Figure 15. While the 'small' end comprises singleton glottal stops, a positivist view could hypothesize that there may be even weaker exponents functioning as filler particles. The demarcation between the 'small' end of the scale and a silent pause is therefore the absence of any interpretable or perceivable phonetic signal. Oriented towards the 'large' end of the continuum are multi-syllabic forms. Around the middle, both non-prototypical and prototypical forms are located. The hypothesis of a continuum of filler particles might also explain multi-segmental phonetic exponents which are impossible to transliterate orthographically, such as the squeaky form observed in Figure 1.

The line of argumentation and the evidence presented concordantly point to the grammatical classification of the phonetic exponents of 'filled pauses' as filler particles, independent of their orthographic representation. As a consequence, this paper argues that grammatical word status can and should be allocated to phonetic exponents even without a conventional orthographic representation. Though it has been argued that filler particles should be classified with interjections (Clark and Fox Tree 2002), their attitudinal as well as distributional features do not support this. However, the approach taken here does support

the claimed word status ([Clark and Fox Tree 2002](#)) of filler particles and even extends this claim beyond prototypical segment sequences. It will be up to future studies to investigate whether glottal filler particles can be classified or interpreted as clitics and which linguistic contexts are preferred by these forms.
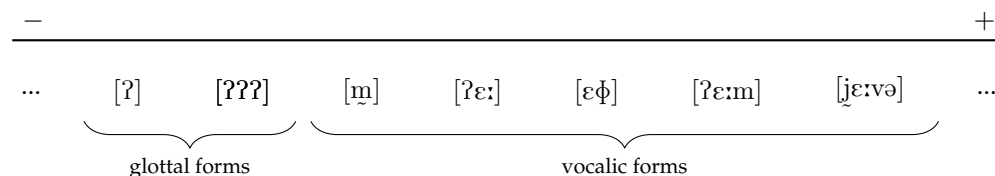
Figure 15. Continuum of phonetic forms of filler particles with German examples from the Berlin Dialogue Corpus.

6.3. Filler/Filler Particle and Their Relevance for Research On (Dis)Fluency

The research on (dis)fluency phenomena is an endeavour rooted in many different disciplines. For any research on (dis)fluencies, especially in linguistics, it is of great value if research can be based on a common terminology and taxonomy, which the term *filler particle* can provide. Of course, even before the categorisation proposed here, there has been a common base of terms that have enabled research on filler particles (e.g., *filled pause* or *filler*). The advantages of adding *particle* to the name of the phenomenon are (at least) twofold. First, the term makes it possible to think about filler particles without focusing on their pragmatic functions, thus overcoming a bias regarding this phenomenon. Secondly, the term anchors the phenomenon in a grammatical category (see Section [6.2](#)). In a further step, the particle can either be recategorized as an interjection, or higher-level pragmatic categories can be assigned to it, such as hesitation marker or similar.

To emphasize, this study is not proposing to rename a phenomenon for the sake of renaming it but rather argues for a more neutral categorization in light of recent research. Categorization, after all, has consequences for how we think about the world. The addition of *particle* to the term *filler* may thus be beneficial for maintaining an unbiased attitude towards them.

Counter to the arguments in favour of the term *filler particle* superseding other terms stands the terminological and taxonomic issue of why the prefix *filler* was chosen, and how the terms relates to *filler particle*. As argued in the introduction, *filler* is a recognisable term in (dis)fluency research which carries almost no semantic baggage. That makes it preferable to, for example, *hesitation particle*. As to the relation of *filler* to *filler particle*, I am not arguing categorically against the use of *filler*. Rather, the aim is to show the advantage of a more precise term that is scientifically and grammatically sound, comes with a definition, and from which future studies may benefit. The term *filler particle* can thus establish a common ground which roots the underlying phenomena in the grammar of a language. This being said, future studies on filler particles are free to insert introductory definitions of the phenomenon after which their usage of the short form *filler* is understood clearly in relation to that. To conclude with a comparison to botany, researchers are still free to decide whether to use the word *berry* or the more precise term *nut* for referring to the fruit of the strawberry plant.

6.4. Implications for Corpus Studies

The introduction of the category of filler particle candidates was motivated by practical considerations, although the discussion above makes it possible to post hoc derive this category from the postulated form continuum of filler particles. Either way, FPCs are practical when annotating filler particles in vast amounts of speech data where it is not feasible to reliably check every detected FPC for its usage, that is, whether the conversational, pragmatic, or prosodic contexts supports the designation of a filler particle. This procedure can, however, constitute a second step after annotating FPCs. Further, FPCs in the signal might not be detected when using automatic transcription or mute corpora.

Although manual annotation of the data is very time-consuming, it seems at this point unavoidable for the reliable detection of glottal filler particles, which add to the frequency of filler particle counts and might—as a speaker-dependent feature—also be useful in forensic phonetics. From a theoretical point of view, the FPC category may also prove to be useful in the discussion of which phonetic exponents have the potential to function as a filler particle.

Both definitions proposed here may hopefully enable researchers to find filler particles which do not match the prototypical phonetic exponents (i.e., form) in a language. This will in turn help to avoid false negatives in quantitative–qualitative corpus studies and to paint a more detailed picture of filler particle frequencies.

6.5. Limitations

There are some limitations with respect to the results of the corpus study. First, not only is the amount of glottal filler particles quite small, there is also a heavy distributional bias towards one of the subcorpora. This might be due to random inter-individual differences, which motivates us both to look at another register with the same individuals and to record another sample with newly sampled individuals. There might also be regional differences in the amount of glottal stops. For example, in GECO (Schweitzer and Lewandowski 2013), a corpus recorded in southern Germany, a study showed that glottal FPs accounted for almost 20% of all filler particles (Belz 2021).

Second, though proposed with the utmost care, it is unclear whether the definition will hold cross-linguistically in typologically different languages. In this regard, Japanese might be a good testing area, as it exhibits phonetic exponents of filler particles which are also similar to lexical items, as in Spanish, but not used in this way. A shortcoming related to this is the lack of a direct comparison of the definition and annotation guideline proposed here to 'conventional' guidelines, which should be examined in future studies, both for German and cross-linguistically.

Third, it remains open for future research to investigate whether listeners actually perceive (directly or indirectly) glottal and vocalic non-prototypical filler particles in the acoustic signal, both as singletons and when presented in context. Related to this is the general question of how many instances transcribers miss when annotating filler particles using the definition presented here, as filler particles are missed more frequently than lexical words (e.g., Le Grezause 2019).

Fourth, filler particle candidates functioning as interjections (such as discussed in Section 5.1.3) were rare in this corpus study, which could be due either to the low frequency of this usage or to the investigated register. The iconic example *um excuse me*, illustrating the interjectional use of filler particles, is expected either when approaching another person unexpectedly (phatic usage) or with an annoyed or disbelieving connotation (expressive usage). Both situations might not be triggered by the register of a conversational dialogue about food, the city of Berlin, or a dream vacation.

7. Conclusions

Although the term *filled pause* (and others) reverberated in linguistics for historical and practical reasons (given that it has been around for over 60 years at least), it comes with some major caveats with respect to its terminological clarity, its set of denotations, and its grammatical classification. This article proposes the term *filler particle* along with a definition to identify its denotations in the acoustic signal and an approach to a grammatical classification. This study showed that filler particles are a well-definable area that can be classified a priori (i.e., before attributing pragmatic functions) within the part-of-speech class of particles and thus can be described within the domain of phonological words. The definition distinguishes filler particles from interjections while being open for higher-level analyses of pragmatic functions in context. Instances of filler particles can be represented on a continuum of forms from singleton segments to multi-syllabic phonetic exponents. Hence, if the reasoning presented here holds, it follows that in the same way chemistry

abandoned the term *phlogiston* after the discovery of *oxygen*, linguistics should dismiss the term *filled pause* after discovering that it is not a pause, but a particle.

Conflicts of Interest: The author declares no conflict of interest.

Notes

[1] In the terminology of corpus linguistics, other terms for phonetic exponents are 'markables'.

[2] It is obvious that this form is only insufficiently covered by *äh*, which is given in lack of a better graphematical representation.

[3] https://www.collinsdictionary.com/dictionary/english/pause, retrieved 20 July 2022.

[4] For the symptom side, cf. Corley and Stewart (2008); for the signal side, cf. Fox Tree and Clark (1997); Siegel et al. (1969); Walker et al. (2014); for a middle way assuming that both hypotheses about filler particles are potentially simultaneously active complementary functions that are not mutually exclusive, cf. Reitbrecht (2017, 27); for a dual status, cf. Kosmala and Crible (2022).

[5] The term *filler* is also used as a short form of *filler sentence* or *filler item* in linguistic experiments, and as a short form of *filler syllable* (referring to protomorphemes) in first language acquisition (Peters 2001). Thus, extending *filler* to *filler particle* comes with the side effect that the polysemous term *filler* can be disambiguated.

[6] Applying Ockham's razor, the characteristics of message delay proposed by Rose (2007) are omitted.

[7] Non-pathological speech is meant in the sense of speech produced by speakers that do not suffer from impairments affecting their speech production, e.g., stuttering.

[8] This does not mean that glottalized V, VN, or N forms are excluded, merely that not every laryngealized segment of a phonological word is automatically judged to be a filler particle.

[9] For example, a cross-linguistic query showed that *äh* in German is listed in dictionaries as equivalent to *esto* in Spanish, *er* in English, and *euh* in French, whereas for *ähm* no results were found (https://tinyurl.com/langenscheidt-aeh, retrieved 13 October 2022).

[10] https://www.dwds.de/wb/gar#d-2-2, retrieved 3 October 2022.

[11] Of course, not knowing the speaker's rationale, we can never rule out a truncation for certain. The context, however, might hint at a specific direction.

[12] https://rs.cms.hu-berlin.de/phon/ retrieved 13 February 2023 .

[13] The subcorpus *videocall* originally included 10 dialogues, of which two apparently included bilingual speakers, which is why they were excluded.

References

Ameka, Felix. 1992. Interjections: The universal yet neglected part of speech. *Journal of Pragmatics* 18: 101–18. [CrossRef]

Ameka, Felix. 2006. Interjections. In *Encyclopedia of Language & Linguistics*. Amsterdam: Elsevier, pp. 743–46. [CrossRef]

Arnold, Jennifer E., Maria Fagnano, and Michael K. Tanenhaus. 2003. Disfluencies Signal Theee, Um, New Information. *Journal of Psycholinguistic Research* 32: 25–36. [CrossRef]

Arnold, Jennifer E., Michael K. Tanenhaus, Rebecca J. Altmann, and Maria Fagnano. 2004. The Old and Thee, uh, New. *Psychological Science* 15: 578–82. [CrossRef] [PubMed]

Bada, Erdoğan, and Bilal Genç. 2008. Pausing preceding and following to in to-infinitives: A study with implications to reading and speaking skills in ELT. *Journal of Pragmatics* 40: 1939–49. [CrossRef]

Baker-Smemoe, Wendy, Dan P. Dewey, Jennifer Bown, and Rob A. Martinsen. 2014. Does Measuring L2 Utterance Fluency Equal Measuring Overall L2 Proficiency? Evidence From Five Languages. *Foreign Language Annals* 47: 707–28. [CrossRef]

Batliner, Anton, Andreas Kießling, Susanne Burger, and Elmar Nöth. 1995. *Filled Pauses in Spontaneous Speech*. Number 88 in Verbmobil Verbundvorhaben.

Belz, Malte. 2017. Glottal filled pauses in German. In *Proceedings of DiSS 2017*. Edited by Robert Eklund and Ralph Rose. TMH-QPSR. Stockholm: KTH Royal Institute of Technology, pp. 5–8.

Belz, Malte. 2020. Acoustic vowel quality of filler particles in German. In *Laughter and Other Non-Verbal Vocalisations Workshop*. Bielefeld: Bielefeld University, pp. 7–10. [CrossRef]

Belz, Malte. 2021. *Die Phonetik von äh und ähm: Akustische Variation von Füllpartikeln im Deutschen*. Berlin: J.B. Metzler. [CrossRef]

Belz, Malte, Alina Zöllner, Megumi Terada, Robert Lange, Lea-Sophie Adam, and Bianca Sell. 2021. Dokumentation und Annotations-richtlinien für das Korpus BeDiaCo v2. Available online: https://doi.org/10.5281/zenodo.4593351 (accessed on 13 February 2023).

Belz, Malte and Jürgen Trouvain. 2019. Are 'silent' pauses always silent? Paper presented at 19th International Congress of Phonetic Sciences (ICPhS), Melbourne, Australia, August 5–9, pp. 2744–48.

Belz, Malte, Christine Mooshammer, Alina Zöllner, and Lea-Sophie Adam. 2021. Berlin Dialogue Corpus (BeDiaCo): Version 2. Available online: https://rs.cms.hu-berlin.de/phon (accessed on 13 February 2023)

Belz, Malte, Simon Sauer, Anke Lüdeling, and Christine Mooshammer. 2017. Fluently disfluent? Pauses and repairs of advanced learners and native speakers of German. *International Journal of Learner Corpus Research* 3: 118–48. [CrossRef]

Betz, Simon. 2020. *Hesitations in Spoken Dialogue Systems*. Bielefeld: Universität Bielefeld. [CrossRef]

Bloomfield, Leonard. 1933. *Language*. London and Aylesbury: Compton.

Boersma, Paul, and David Weenink. 2022. *Praat: Doing Phonetics by Computer*, version 6.3. Available online: http://www.praat.org (accessed on 13 February 2023).

Boomer, Donald S., and Allen T. Dittmann. 1962. Hesitation pauses and juncture pauses in speech. *Language and Speech* 5: 215–20. [CrossRef]

Brown, Keith, and Jim Miller. 2013. *The Cambridge Dictionary of Linguistics*. Cambridge and New York: Cambridge University Press. [CrossRef]

Chafe, Wallace L. 1980. Some reasons for hesitating. In *Temporal variables in Speech*. Edited by Hans W. Dechert and Manfred Raupach. The Hague: Mouton, pp. 169–82.

Clark, Herbert H., and Jean E. Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition* 84: 73–111. [CrossRef] [PubMed]

Collard, Philip, Martin Corley, Lucy J. MacGregor, and David I. Donaldson. 2008. Attention orienting effects of hesitations in speech: Evidence from ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34: 696–702. [CrossRef]

Corley, Martin, and Oliver W. Stewart. 2008. Hesitation Disfluencies in Spontaneous Speech: The Meaning of um. *Language and Linguistics Compass* 2: 589–602. [CrossRef]

Crible, Ludivine. 2017. Towards an operational category of discourse markers: A definition and its model. In *Pragmatic Markers, Discourse Markers and Modal Particles*. Edited by Fedriani Chiara and Sansò Andrea. Studies in Language Companion Series. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 99–124.

Crible, Ludivine, Liesbeth Degand, and Gaëtanelle Gilquin. 2017. The Clustering of Discourse Markers and Filled Pauses. *Languages in Contrast: International Journal for Contrastive Linguistics* 17: 69–95. [CrossRef]

Crystal, David. 2008. *Dictionary of Linguistics and Phonetics*, 6th ed. Hoboken: John Wiley & Sons Incorporated.

de Boer, Meike M., Hugo Quené, and Willemijn F. L. Heeren. 2022. Long-term within-speaker consistency of filled pauses in native and non-native speech. *JASA Express Letters* 2: 035201. [CrossRef] [PubMed]

de Leeuw, Esther. 2007. Hesitation Markers in English, German, and Dutch. *Journal of Germanic Linguistics* 19: 85–114. [CrossRef]

Dingemanse, Mark. 2020. Between Sound and Speech: Liminal Signs in Interaction. *Research on Language & Social Interaction* 53: 188–96. [CrossRef]

Ehlich, Konrad. 1986. *Interjektionen*. Volume 111 of *Linguistische Arbeiten*. Tübingen: Max Niemeyer Verlag.

Eklund, Robert. 2004. Disfluency in Swedish Human-Human and Human-Machine Travel Booking Dialogues. Ph.D. thesis, Linköpings Universitet, Linköping, Sweden.

Finlayson, Ian R., and Martin Corley. 2012. Disfluency in dialogue: An intentional signal from the speaker? *Psychonomic Bulletin & Review* 19: 921–8. [CrossRef]

Fischer, Kerstin. 2006. Towards an understanding of the spectrum of approaches to discourse particles: Introduction to the volume. In *Approaches to Discourse Particles*. Edited by Kerstin Fischer. Studies in Pragmatics. Leiden: Brill, pp. 1–20.

Fox Tree, Jean E., and Herbert H. Clark. 1997. Pronouncing "the" as "thee" to signal problems in speaking. *Cognition* 62: 151–67. [CrossRef]

Fraundorf, Scott H. and Duane G. Watson. 2013. Alice's adventures in um-derland: Psycholinguistic sources of variation in disfluency production. *Language and Cognitive Processes* 29: 1083–96. [CrossRef]

Gick, Bryan, Ian Wilson, Karsten Koch, and Clare Cook. 2004. Language-Specific Articulatory Settings: Evidence from Inter-Utterance Rest Position. *Phonetica* 61: 220–33. [CrossRef]

Gilquin, Gaëtanelle. 2008. Hesitation markers among EFL learners: Pragmatic deficiency or difference? In *Pragmatics and Corpus Linguistics: A Mutualistic Entente*. Edited by Jesus Romero-Trillo. Mouton Series in Pragmatics (MSP). Berlin, Heidelberg and New York: Mouton de Gruyter, pp. 119–49.

Gläser, Rosemarie. 1995. Eigennamen in Wissenschafts- und Techniksprache. In *Namenforschung/Name Studies/Les noms propres: 1. Halbband*. Edited by Ernst Eichler, Gerold Hilty, Heinrich Loffler, Ladislav Zgusta and Hugo Steger. Handbücher zur Sprach- und Kommunikationswissenschaft. Berlin and New York: Walter de Gruyter, pp. 527–33.. [CrossRef]

Goldman-Eisler, Frieda. 1961. A comparative study of two hesitation phenomena. *Language and Speech* 4: 18–26. [CrossRef]

Götz, Sandra. 2019. Filled pauses across proficiency levels, L1s and learning context variables: A multivariate exploration of the Trinity Lancaster Corpus Sample. *International Journal of Learner Corpus Research* 5: 159–80. [CrossRef]

Graham, Lamar A. 2018. Variation in hesitation. *Spanish in Context* 15: 1–26. [CrossRef]

Hawkins, P. R. 1971. The syntactic location of hesitation pauses. *Language and Speech* 14: 277–88. [CrossRef] [PubMed]

Hoey, Elliott M. 2020. Waiting to Inhale: On Sniffing in Conversation. *Research on Language & Social Interaction* 53: 118–39. [CrossRef]

Jessen, Michael. 2012. *Phonetische und linguistische Prinzipien des forensischen Stimmenvergleichs*. Volume 9 of *LINCOM Studies in Phonetics*. München: LINCOM.

Jucker, Andreas H. 2015. Uh and um as planners in the Corpus of Historical American English. In *Developments in English*. Edited by Irma Taavitsainen, Merja Kytö, Claudia Claridge and Jeremy Smith. Studies in English language. Cambridge: Cambridge University Press, pp. 162–77. [CrossRef]

Keseling, Gisbert. 1989. Die Partikel ÄH. Ein paraverbales Element im Sprachsystem? In *Sprechen mit Partikeln*. Edited by H. Weydt. Berlin and New York: Walter de Gruyter, pp. 575–91.

Kirjavainen, Minna, and Alexandre Nikolaev. 2022. Investigation into the linguistic category membership of the finnish planning particle 'tota'. *Pragmatics & Cognition* 29: 375–98. [CrossRef]

Kirjavainen, Minna, Ludivine Crible, and Kate Beeching. 2022. Can filled pauses be represented as linguistic items? Investigating the effect of exposure on the perception and production of um. *Language and Speech* 65: 263–89. [CrossRef]

Kjellmer, Göran. 2003. Hesitation. In Defence of ER and ERM. *English Studies* 84: 170–98. [CrossRef]

Klug, Katharina. 2013. 'Ähm'—Sind Häsitations-Partikeln sprecherspezifisch? Untersuchung der Parameter Grundfrequenz und Vokalqualität. In *Aktuelle Forschungsthemen der Sprechwissenschaft 3*. Edited by von Lutz Christian Anders, Ines Bose, Ursula Hirschfeld and Baldur Neuber. Hallesche Schriften zur Sprechwissenschaft und Phonetik (HSSP). Frankfurt am Main: Peter Lang, pp. 65–94.

Kohler, Klaus J., Benno Peters, and Thomas Wesener. 2005. Phonetic Exponents of Disfluency in German Spontaneous Speech. In *Prosodic Structures in German Spontaneous Speech*. Kiel: Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel, pp. 185–201.

Kosmala, Loulou, and Ludivine Crible. 2022. The dual status of filled pauses: Evidence from genre, proficiency and co-occurrence. *Language and Speech* 65: 216–39. [CrossRef]

Kowal, Sabine, and Daniel C. O'Connell. 2004. Interjektionen im Gespräch. *Zeitschrift für Semiotik* 26: 85–100.

Künzel, Hermann J. 1987. *Sprechererkennung: Grundzüge forensischer Sprachverarbeitung*. Heidelberg: Kriminalistik Verlag.

Le Grezause, Esther. 2019. Um and Uh, and the Expression of Stance in Conversational Speech. Ph.D thesis, Université de Sorbonne Paric Cité and University of Washington, Washington, DC, USA.

Li, Xiaoting. 2020. Click-Initiated Self-Repair in Changing the Sequential Trajectory of Actions-in-Progress. *Research on Language & Social Interaction* 53: 90–117. [CrossRef]

Lickley, Robin J. 2001. Dialogue moves and disfluency rates. Paper presented at ITRW on Disfluency in Spontaneous Speech (DiSS'01), Edinburgh, UK, August 29–31. pp. 93–96.

Lickley, Robin J. 2015. Fluency and Disfluency. In *The Handbook of Speech Production*. Edited by Melissa A. Redford. Hoboken: John Wiley & Sons, Inc., pp. 445–69. [CrossRef]

Linell, Per. 1982. *The Written Language Bias in Linguistics*. Volume 2 of *Studies in Communication*. Linköping: University of Linköping.

Maclay, Howard and Charles E. Osgood. 1959. Hesitation Phenomena in Spontaneous English Speech. *Word* 5: 19–44. [CrossRef]

Mugdan, Joachim. 1990. On the History of Linguistic Terminology. In *History and Historiography of Linguistics*. Edited by Hans-Josef Niederehe and E. F. K. Koerner. Amsterdam studies in the theory and history of linguistic science. Amsterdam and Philadelphia: John Benjamins, pp. 49–61.

Nicholson, Hannele Buffy Mary. 2007. Disfluency in Dialogue: Attention, Structure and Function. Ph.D. thesis, University of Edinburgh, Edinburgh, UK.

Nicholson, Hannele Buffy Mary, Kathleen Eberhard, and Matthias Scheutz. 2010. "Um... I don't see any": The Function of Filled Pauses and Repairs. Paper presented at DiSS-LPSS Joint Workshop 2010, The 5th Workshop on Disfluency in Spontaneous Speech and The 2nd International Symposium on Linguistic Patterns in Spontaneous Speech, Tokyo, Japan, September 25–26.

O'Connell, Daniel C. and Sabine Kowal. 2005. Uh and Um Revisited: Are They Interjections for Signaling Delay? *Journal of Psycholinguistic Research* 34: 555–76. [CrossRef] [PubMed]

Ogden, Richard. 2020. Audibly Not Saying Something with Clicks. *Research on Language & Social Interaction* 53: 66–89. [CrossRef]

Peters, Ann M. 2001. Filler syllables: What is their status in emerging grammar? *Journal of Child Language* 28: 229–42. [CrossRef] [PubMed]

Peters, Benno. 2005. Weiterführende Untersuchungen zu prosodischen Grenzen in deutscher Spontansprache. In *AIPUK 35a*. Kiel: Universität Kiel, pp. 203–345.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing (v. 4.1.3)*. Vienna: R Foundation for Statistical Computing.

Reitbrecht, Sandra. 2017. *Häsitationsphänomene in der Fremdsprache Deutsch und ihre Bedeutung für die Sprechwirkung*. Volume 10 of *Schriften zur Sprechwissenschaft und Phonetik*. Berlin: Frank & Timme.

Rendle-Short, Johanna. 2004. Showing structure: Using um in the academic seminar. *Pragmatics* 14: 479–98. [CrossRef]

Riggenbach, Heidi. 1991. Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes* 14: 423–41. [CrossRef]

Roggia, Aaron B. 2012. Eh as a polyfunctional discourse marker in Dominican Spanish. *Journal of Pragmatics* 44: 1783–98. [CrossRef]

Rose, Ralph L. 2007. What Is a Filled Pause? Available online: http://filledpause.com/musings/what-filled-pause (accessed on 13 February 2023).

Rose, Ralph L. 2015. Um and uh as differential delay markers: The role of contextual factors. Paper presented at Disfluency in Spontaneous Speech (DiSS), Scotland, UK, August 8–9. pp. 73–76.

Schegloff, Emanuel A. 2010. Some Other "Uh(m)"s. *Discourse Processes* 47: 130–74. [CrossRef]

Schneider, Ulrike. 2014. Frequency, Chunks and Hesitations: A Usage-Based Analysis of Chunking in Englisch. Ph.D. thesis, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany. [CrossRef]

Schweitzer, Antje, and Natalie Lewandowski. 2013. Convergence of Articulation Rate in Spontaneous Speech. Paper presented at Interspeech 2013, Lyon, France, August 25–29. pp. 525–9.

Shriberg, Elizabeth E. 1994. Preliminaries to a Theory of Speech Disfluencies. Unveröffentlichte dissertation, University of California, Berkeley, CA, USA.

Shriberg, Elizabeth E., and Robin J. Lickley. 1993. Intonation of clause-internal filled pauses. *Phonetica* 50: 172–9. [CrossRef]

Siegel, Gerald M., Joanne Lenske, and Patricia Broen. 1969. Suppression of normal speech disfluencies through response cost. *Journal of Applied Behavior Analysis* 2: 265–276. [CrossRef] [PubMed]

Smith, Vicki L., and Herbert H. Clark. 1993. On the Course of Answering Questions. *Journal of Memory and Language* 32: 25–38. [CrossRef]

Swerts, Marc. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30: 485–96. [CrossRef]

Tottie, Gunnel. 2015. Turn management and the fillers uh and uhm. In *Corpus Pragmatics*. Edited by Karin Aijmer and Christoph Ruhlemann. Cambridge: Cambridge University Press, pp. 381–407.

Tottie, Gunnel. 2016. Planning what to say: Uh and um among the pragmatic markers. In*Outside the Clause*. Edited by Gunther Kaltenböck, Evelien Keizer and Arne Lohmann. Volume 178 of *Studies in Language Companion Series*. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 97–122. [CrossRef]

Trouvain, Jürgen. 2014. Laughing, breathing, clicking—The prosody of nonverbal vocalisations. In *Speech Prosody*. pp. 598–602. Dublin: Trinity College Dublin.

Trouvain, Jürgen. 2015. On clicks in German. In *Trends in Phonetics and Phonology*. Edited by Adrian Leemann, Marie-José Kolly, Stephan Schmid and Volker Dellwo. Bern: Peter Lang, pp. 21–34.

Trouvain, Jürgen and Raphael Werner. 2022. A phonetic view on annotating speech pauses and pause-internal phonetic particles. In *Transkription und Annotation gesprochener Sprache und multimodaler Interaktion*. Edited by Grawunder Sven and Schwarze Cordula. Tübingen: Narr, pp. 55–73.

Walker, Esther J., Evan F. Risko, and Alan Kingstone. 2014. Fillers as Signals: Evidence From a Question–Answering Paradigm. *Discourse Processes* 51: 264–86. [CrossRef]

Watanabe, Michiko, Keikichi Hirose, Yasuharu Den, and Nobuaki Minematsu. 2008. Filled pauses as cues to the complexity of upcoming phrases for native and non-native listeners. *Speech Communication* 50: 81–94. [CrossRef]

Wieling, Martijn, Jack Grieve, Gosse Bouma, Josef Fruehwald, John Coleman, and Mark Liberman. 2016. Variation and change in the use of hesitation markers in Germanic languages. *Language Dynamics and Change* 6: 199–234. [CrossRef]

Willkop, Eva-Maria. 1988. *Gliederungspartikeln im Dialog*. Volume 5 of *Studien Deutsch*. München: Iudicium.

Winkelmann, Raphael, Jonathan Harrington, and Klaus Jänsch. 2017. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language* 392–410. [CrossRef]

Winkelmann, Raphael, Klaus Jaensch, Steve Cassidy, and Jonathan Harrington. 2020. emuR: Main Package of the EMU Speech Database Management System. Available online: https://github.com/IPS-LMU/emuR (accessed on 13 February 2023).