The acoustic features and didactic function of Foreigner Directed Speech: A scoping review.

Giorgio Piazza[a,b]*, Clara D. Martin[a,c], Marina Kalashnikova[a,c]

[a]*Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastián, Spain;*

[b]*Department of Social Sciences and Law, UPV/EHU, Donostia- San Sebastián, Spain;*

[c]*Ikerbasque, Basque Foundation for Science, Bilbao, Spain.*

*Corresponding author: g.piazza@bcbl.eu; Mikeletegi Pasealekua, 69, 20009 Donostia-San Sebastián, Gipuzkoa, Spain.

This is the ACCEPTED MANUSCRIPT of the paper in press on the Journal of Speech, Language and Hearing Research. https://pubs.asha.org/journal/jslhr

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that would influence the work reported in this paper.

Funding sources

This research was supported by a Doctoral Fellowship (LCF/BQ/DI19/11730045) from "La Caixa" Foundation (ID 100010434) to G.P., and by the Spanish Ministry of Science and Innovation through the Ramon y Cajal Research Fellowship (RYC2018-024284-I) to M.K. This research was supported by the Basque Government through the BERC 2022-2025 program and by the Spanish State Research Agency through BCBL Severo

Ochoa excellence accreditation CEX2020-001010-S. The research was also supported by the Spanish Ministry of Economy and Competitiveness (PID2020-113926GB-I00 to C.D.M.), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 819093 to C.D.M.).

Abstract

Purpose. This scoping review considers the acoustic features of a clear speech register directed to non-native listeners known as Foreigner Directed Speech (FDS). We identify vowel hyperarticulation and low speech rate as the most representative acoustic features of FDS; other features, including wide pitch range and high intensity, are still under debate.

We also discuss factors that may influence the outcomes and characteristics of FDS. We start by examining accommodation theories, outlining the reasons why FDS is likely to serve a didactic function by helping listeners acquire a second language (L2). We examine how this speech register adapts to listeners' identities and linguistic needs, suggesting that FDS also takes listeners' L2 proficiency into account. To confirm the didactic function of FDS, we compare it to other clear speech registers, specifically Infant Directed Speech and Lombard Speech.

Conclusion. Our review reveals that research has not yet established whether FDS succeeds as a didactic tool that supports L2 acquisition. Moreover, a complex set of factors determine specific realizations of FDS, which need further exploration. We conclude by summarising open questions and indicating directions and recommendations for future research.

Keywords

Foreigner-directed speech; listener-oriented speech; speech accommodation; didactic function; speech registers; second language learning.

The acoustic features and didactic function of Foreigner Directed Speech: A scoping review.

1. What is Foreigner Directed Speech?

Foreigner Directed Speech (henceforth, FDS) is a speech register that native speakers use in interactions with non-native speakers of their language. In recent literature, this register has also been referred to as “L2 speech accommodation”. FDS is a broad phenomenon that can encompass changes at the discourse, syntactic, lexical, and acoustic levels (Chaudron, 1979; Long, 1981; Ramamurti, 1980; Uther et al., 2007). FDS is proposed to be a – mostly unconscious – speech accommodation that increases speech clarity and that could serve a didactic function by assisting non-native interlocutors to better understand, perceive, and articulate their L2 (Hatch, 1979; Tarone, 1980; Uther et al., 2007; Scarborough et al., 2007). In this article, we provide a critical scoping review of the extensive research investigating the didactic function proposal, focusing on the acoustic features of FDS. We propose that the *didactic function* of FDS comprises two related aspects: a didactic purpose, which is the function of teaching an L2, reflected on the acoustic features of FDS, and a didactic impact, which is the actual effect on L2 perception and learning. In light of this didactic function, we discuss whether FDS serves a purpose in increasing speech intelligibility, facilitating L2 learning, and whether L2 listeners may benefit from being exposed to FDS. The objective of this work is to review those aspects of FDS that are still under debate and to provide strong theoretical and methodological bases for future research into this speech register. An in-depth study of the features and function of FDS is expected to increase our understanding of communication between humans who do not share the same native language. This will enable researchers to build appropriate models of speech communication and social mediation. As linguistic diversity increases worldwide, making communication between native and non-native speakers ever more frequent, speech communication models need to account for FDS.

One of the earliest mentions of FDS as a speech register that serves a linguistic function is found in the 1930’s, when Bloomfield (1933) proposed that FDS reflects native speakers’ tendency to imitate the mistakes made by non-native speakers in order to assist their speech comprehension. Several decades later, FDS was positioned as a variant of *clear speech*, a term typically used to refer to registers that enhance speech clarity. Other clear speech variants include Infant Directed Speech (IDS, also known as baby talk) and speech directed to elders or to people with hearing impairments (Ferguson, 1975; Lam et al., 2012; see also Smiljanić & Bradlow, 2009 for a review on clear speech). It was only in the 1970's that FDS was identified as an independent speech register that – despite sharing some features with other registers (such as IDS, see Section 6.1) – manifests in speech specifically directed to non-native listeners and is uniquely suited to their linguistic needs (Ferguson, 1975). Ferguson (1971, 1975) coined the term *foreign talk* to implicitly compare this register to baby talk, suggesting that the two registers shared a didactic function (Hatch, 1979; Katz, 1977; Tarone, 1980; see also Kuhl et al., 1997 for a discussion on the didactic functions of IDS). FDS was proposed to convey articulatory instructions through a simplified register (as Ferguson, 1981 defined it) characterized by repetition and the use of high frequency words, reduced syntactic complexity, and lack of jargon or idiomatic expressions (Chaudron, 1979; Long, 1981, 1983; Ramamurti, 1980). Additional features, also assumed to assist L2 learners, have been proposed in contemporary studies, including low speech rate, exaggerated voicing of final stops, few vowel reductions, as well as exaggerated intonation and volume (Hatch et al., 1978; Hatch, 1979 reported by Tarone, 1980).

At present, there is more extensive knowledge of FDS, and research largely focuses on the acoustic features of this register. It has been shown that various acoustic features are the result of the FDS accommodation: vowel hyperarticulation, low speech rate, and long pauses are all proposed to help non-native listeners. Given that interest in this topic is growing, there is a need for a review that sums up and discusses the most relevant findings on FDS and sets goals for future research on this topic. In the past, research on FDS has focused on defining the acoustic features of this register; although some FDS features continue to be the subject of debate, future studies should focus on advancing our understanding of the factors that underlie these FDS adjustments, and the role that each FDS feature plays in non-native listeners' L2 acquisition. The present paper provides the starting point for addressing these issues. This review includes all journal articles and conference proceedings available to date that (1) were written in English and (2) reported empirical findings. These were identified by an extensive literature search using the Google Scholar search engine (search terms: FDS, Foreigner directed speech, Foreign Talk, non-native directed speech, speech accommodation, listener-oriented speech) and complemented by including relevant references from the articles.

Here, we discuss important aspects of FDS research that help to discover its role in the native–non-native interaction. In Section 2, we focus on the acoustic features of FDS to explain *how* FDS improves speech clarity; in Section 3 we discuss the emotional valence of FDS that differs for native and non-native listeners. Section 4 frames the accommodation theories behind FDS, whereas Section 5 and 6 discuss whether FDS is adjusted to the listener's needs and thus supports L2 acquisition. Section 7 describes research on native and non-native listeners' perception of FDS – which will help us understand whether FDS is useful to non-native listeners. Section 8 presents our conclusions and recommendations for future research on this topic.

2. The acoustic features of FDS.

2.1 Vowel hyperarticulation

Compared to Native Directed Speech (NDS), the register used by peers sharing the same native language who have no need to further enhance intelligibility (Ferguson & Kewley-Port, 2002; Smiljanić & Bradlow, 2009), FDS is characterized by an expanded vocalic space that is known as vowel hyperarticulation (Uther et al., 2007; Scarborough et al., 2007; Knoll et al., 2007; Knoll et al., 2009a). Most studies on FDS vowel hyperarticulation focus on native speakers' production of the three corner vowels /a/, /i/, and /u/. These vowels are considered important because they are located at the outer boundaries (low, frontal, and posterior) of a language's vocalic space, and they are present in the vocalic inventories of most languages in the world (Bradlow et al., 2003; Ladefoged & Maddieson, 1990). Usually, the averages of the first (F1) and second (F2) formant values are projected onto a two-dimensional cartesian plane to form the corners of the vocalic triangle. An expanded vocalic space corresponds to a vocalic triangle with a larger area (see Figure 1) since the corner vowels are produced at a greater distance from each other. This is proposed to enhance detection of vocalic contrasts and aid speech perception and comprehension (Bradlow & Bent, 2002; Ferguson & Kewley-Port, 2002; Smiljanić & Bradlow, 2005, 2009). F1 inversely relates to vowel height (the lower the vowel articulation, the higher its F1 value) and reflects articulatory effort, which is commonly higher in all clear speech styles than in conversational speech (Ladefoged, 2006; Smiljanić and Bradlow, 2009). F2 is affected by posteriority and lip rounding (F2 values are lower for vowels produced further back in the vocal tract; Ladefoged, 2006). F2 height is usually influenced by vowel hyperarticulation, but it also depends on whether speakers are expanding a front vowel (higher F2) or a back vowel (lower F2) (Ferguson & Kewley-Port, 2002).

INSERT FIGURE 1 ABOUT HERE

Of particular relevance to this review, vowel hyperarticulation is proposed to be the key acoustic feature that serves a didactic function (both purpose and impact) because it results in a clearer and more distinctive representation of vowel categories (Kuhl et al., 1997). The expansion of the vowel triangle allows speakers to create more discrete categories, thereby avoiding confusion and overlap between vowels and supporting vowel imitation and feature acquisition (Kuhl et al., 1997). Further supporting its proposed didactic purpose, vowel hyperarticulation occurs across clear registers associated with higher speech intelligibility (Bradlow et al., 2003; Krause & Braida, 2004; Picheny et al., 1986), but it is restricted to registers directed to audiences with perceived linguistic capacity (Burnham et al., 2002). For instance, vowel hyperarticulation has been reported in speech to infants (Kuhl et al., 1997) and to computer avatars (Burnham et al., 2010), but not in speech to pets such as cats and dogs (Burnham et al., 2002) – unless the pet is a parrot that “repeats” words (Xu et al., 2013). Thus, vowel hyperarticulation might also be expected in FDS, since speakers’ production is presumably modulated to support the fledgling linguistic abilities of the L2 learner.

We were able to identify 12 studies published to date, that have investigated the presence of vowel hyperarticulation in FDS: 8 studies out of a total of 12 reported vowel hyperarticulation in FDS. Most studies focused on English FDS and identified vowel hyperarticulation as the main feature differentiating

FDS from NDS (Knoll et al., 2007; 2009a; Uther et al., 2007; Scarborough et al., 2007; Hazan et al., 2015). At the best of our knowledge, only Kendi and Khattab (2019) have reported vowel hyperarticulation in a language other than English, providing some evidence for cross-linguistic generalization. However, while some previous research had reported both F1 and F2 exaggeration, Kendi and Khattab (2019) found Arabic (Omani) speakers only used F1 to expand their vowel space (consistent with research by Dodane & Al-Tamimi, 2007). This is interesting in light of literature that claims that the degree of vowel hyperarticulation (in clear speech in general) depends on the vowel inventory size of each language. This would suggest that hyperarticulation is language-dependent, and likely to be predominant in languages with large vowel inventories (e.g., Andruski et al., 1999; see also Al-Tamami & Ferragne, 2005). Despite the relatively small number of vowels in Arabic (6 vowels as compared to 14 in English) (Saadah, 2011), Kendi and Khattab (2019) confirmed the presence of vowel hyperarticulation in Arabic FDS. It could be that modulation of hyperarticulation is not fully determined (Smiljanić and Bradlow, 2005) but instead varies in relation to the size of vowel inventories.

Despite the large number of studies confirming that vowel hyperarticulation is present in FDS, some studies have reported different result patterns even for FDS produced in English (Knoll & Scharrer, 2007; Knoll et al., 2009a, 2011a; Kangatharan et al., 2012). For instance, Kangatharan et al. (2012) indicated that they failed to find vowel hyperarticulation in English FDS. This may be because their assessment relied on a vocalic square (instead of triangle) that included vowels /e/, /iː/, /ɔː/ and the diphthong /ai/, which are usually not considered in these kinds of studies. Some of those mixed results derive from specific manipulations of the experimental design (e.g., use of imaginary listeners, see Section 5.1, see also Knoll & Scharrer, 2007; Knoll et al., 2009a). Despite such conflicting evidence, given the findings reported above, we conclude that vowel hyperarticulation is a robust feature of FDS. Nevertheless, it remains unclear how this feature relates to the vowel inventory size of a given language. Further cross-linguistic investigations including languages other than English will be needed to determine whether vowel hyperarticulation varies across languages (language-specificity).

In addition to adjustments to vowel formants, vowel hyperarticulation may also be manifested as vowel lengthening. Longer vowels may be easier to process and categorize, vowel lengthening thus having a didactic impact (see Biersack et al., 2005). In order to explore the potential relevance of vowel lengthening in FDS, Uther et al. (2007) analysed whether vowel hyperarticulation of FDS (and IDS) in their study was also associated with longer vowel durations. Instead, they found that vowel length in FDS did not differ from NDS (and was shorter than in IDS). This pattern has been confirmed in other studies where FDS vowel length was not associated with the expansion of vowel space (Knoll & Scharrer, 2007; Knoll et al., 2009a, 2011a; cf. Biersack 2005). Thus, it appears that hyperarticulation of vowels is a clear acoustic feature associated with FDS, but that these vowels are not produced with longer duration. This confirms the proposal from Biersack et al. (2005) that FDS should give non-native listeners more time to process sentences by lengthening pauses, but not vowels.

Some attention has also been dedicated to hyperarticulation of non-vowel phonological contrasts. For instance, Sankowska et al. (2011) were the first to find that the plosive durational difference (between voiced and voiceless consonants) was larger in FDS than in either NDS or Lombard Speech (LS), a register produced to help listeners cope with background noise. This finding suggests that hyperarticulation in FDS might not be relegated to vowel articulation only. Similarly, several studies have reported hyperarticulation of lexical tone categories in Foreigner (Papoušek & Hwang, 1991) and Infant Directed Speech (Han et al., 2018; Liu et al., 2007; Xu et al., 2013) in lexical tone languages. Lexical tones are based primarily on modulations of pitch height and contour, and their realization is not independent of segments (tones are carried by vowels as well as the adjacent consonants, so they can be considered supra-segmental; Burnham et al., 2011). However, similarly to consonant and vowel segments, lexical tones mark phonemic contrasts, and so it is not surprising that lexical tone categories are also exaggerated in clear speech registers.

2.2 Low speech rate and long pauses

Another characteristic acoustic feature of FDS is low speech rate, measured as an increase in pauses between utterances and the duration of individual words within utterances (Ferguson, 1975; Biersack et al., 2005; Scarborough et al., 2007; Kangatharan, 2015; Lorge & Katsos, 2019; Bobb et al., 2019).

FDS tends to have a lower rate of words per minute than NDS (Biersack et al., 2005; Hatch, 1979; Nelson, 1992; Rodriguez-Cuadrado et al., 2018; Scarborough et al., 2007). Similar to vowel hyperarticulation, this phenomenon is proposed to benefit comprehension and processing of speech by non-native listeners: L2 listeners might benefit from having more time to parse, segment, and analyse linguistic information when speech rate is slower (see Biersack et al., 2005).

While vowel hyperarticulation across languages awaits further research, speech rate adjustments in FDS have been explored directly in cross-linguistic studies (e.g., English vs. French). Warren-Leubecker & Bohannon (1982) and Hazan et al. (2015) reported that English FDS has a lower word rate per minute than NDS. Kühnert and Antolík (2017) provided evidence from French using a *tandem* paradigm: participants with two different native languages (L1s) practiced language exchange to help each other learn an L2; each participant was a native speaker of the L2 the other participant wished to learn. Using this paradigm in English and French, Kühnert and Antolík (2017) found that French native speakers accommodated their production to the English listeners (French L2 learners) by slowing down their speech rate. However, native English speakers did not significantly lower their speech rate when they interacted with the French (English L2). This finding on English speakers contrasts both with the results for the French participants in this study and with previous studies on English FDS. The authors explained this incongruency by suggesting that speech tempo adjustments may be language specific, possibly related to the faster natural speech rate in French than English, or to French participants having higher proficiency in English than their English counterparts had in French. In fact, slower speech was directed to L2 listeners with lower proficiency (English native speakers), whereas the faster speech rate was directed to highly proficient L2 listeners (French native speakers). This suggests that low speech rate may be a feature of FDS directed to naïve L2 learners in order to support their L2 comprehension.

2.3 Intensity and Pitch

Several suprasegmental features including acoustic intensity, pitch height, pitch range, and pitch contours have also been investigated in FDS. It remains unclear whether enhancement or exaggeration of these features serve independent didactic purposes or occur as by-products of the phonetic exaggeration already noted in this register. For instance, vowel hyperarticulation may be accompanied by increased intensity and heightened pitch whereas the independent exaggeration of either of these two features may not result in enhanced clarity.

Intensity corresponds to the amount of energy carried by sound waves, and loudness is its primary perceptual correlate. High intensity is a prosodic cue for emphasis that in FDS might correlate with vowel hyperarticulation. Vowel hyperarticulation might be the result of articulatory effort (see Lindblom, 1990), while intensity and loudness are secondary correlates of effort (see also Smiljanić and Bradlow, 2009). Rodriguez-Cuadrado et al. (2018) analysed Spanish FDS by measuring the intensity of repeated words, which are usually hypoarticulated in conversational speech. They observed higher word intensity for repeated words in FDS than NDS. Kendi and Khattab (2019), in their study on Arabic FDS, also reported significantly higher vowel intensity in FDS than NDS (in line with Hazan et al., 2015; cf. Knoll et al., 2015). Thus, higher intensity could be a relevant feature of FDS, but the evidence for this claim to date is not robust, and no studies have directly tested whether it is an acoustic correlate of another FDS feature, specifically vowel hyperarticulation (see Ferguson & Quené, 2014 for similar a hypothesis regarding clear speech).

Pitch is a suprasegmental feature that is used to mark prosody conveying prominence and/or information structure. Pitch is the perceived acoustic product of the vibration rate of the vocal cords (Ladefoged, 2006). Pitch range corresponds to the maximal/minimal excursions of pitch (i.e., the difference between pitch *Max* and *Min*), whereas pitch contour is the curve of the perceived pitch change over time. Very few studies have investigated pitch range in FDS, and existing studies have reported both wider pitch excursions compared to NDS (Smith, 2007) and comparable pitch ranges for the two registers (Knoll et al., 2015). It would not be surprising to find an expanded pitch range in FDS since emphatic pitch excursion is proposed to form part of hyperarticulatory phenomena, stressing relevant words and assisting word segmentation (as research on IDS suggests: Fernald & Kuhl, 1987; Thiessen et al., 2005).

Although exaggerated pitch contour is linked to pitch range, the two features are not equivalent. Specifically, the same pitch range value could be associated with both bell and rising contour shapes and vice versa. In fact, despite some evidence for a wider pitch range in FDS than NDS (Papoušek & Hwang, 1991; Smith, 2007), the few experiments that assessed pitch contour reported little evidence for exaggerated pitch contours in FDS (Papoušek & Hwang, 1991; Knoll et al., 2006, 2007; Knoll & Costall, 2015). Knoll et al. (2006) studied contour shape and found that FDS did not contain exaggerated shapes compared to NDS (Knoll et al., 2007). Most contours in FDS were level shapes (flat) or falling contours like those found in NDS (Knoll & Costall, 2015). Further qualitative analysis in Knoll and Costall (2015) highlighted the fact that the participants (students) sometimes produced FDS with rising contours, which are similar to the contour shape of questions (Fernald, 1989; Knoll et al., 2006, 2007; Knoll & Costall, 2015). However, it is likely that native speakers were trying to assess the L2 listener's comprehension so as to adapt their production accordingly and used a rising contour typically associated with a questioning tone (a silent "did you understand?") to implicitly interrogate listeners' comprehension. This strategy invites the listener to provide continuous feedback, either verbally or nonverbally (e.g., through nods, confused expressions; see section 5.3 for a discussion of listeners' feedback). This interpretation of the rising contour converges with the reported results on pitch range in IDS, which show mostly bell shape contours. The main hypothesis is that exaggerated pitch contour and wide pitch excursion serve the functions of emotional transfer and requesting attention (Ferguson, 1971; Papoušek & Papoušek, 1981; Trainor & Desjardins, 2002). The results on FDS fit with this view since speakers are not expected to employ FDS to convey emotions but rather to draw the listener's attention to meaningful words.

Mean pitch exaggeration corresponds to raised fundamental frequency (F_0), which is found in other clear speech registers such as IDS. Exaggerated mean pitch is mostly interpreted as a strategy to convey emotion and a non-threatening attitude in speech (Ohala, 1984), and for this reason this feature is not expected to be prominent in FDS. In fact, the many studies sustain that FDS does not have a high pitch correlate (Biersack et al., 2005; Bobb et al., 2019; Knoll et al., 2009a, 2011a; 2011b; Lorge & Katsos, 2019; Uther et al., 2007). For example, Biersack et al. (2005) and Uther et al. (2007) found no significant difference between FDS and NDS. It is important to note that although some studies underscored the absence of heightened mean pitch in FDS, other studies have disclosed a different pattern of results. In recent work, Kendi and Khattab (2019) demonstrated that pitch average midpoints were higher in FDS than NDS (Hazan et al., 2015). However, this significant increase in mean pitch in FDS might be due to specific aspects of their study design. In Kendi and Khattab (2019), the addressees of FDS were domestic helpers who had been working for the participants (i.e., speakers) for an extended period of time (from 2 months to 4 years). Indeed, previous evidence has shown that familiarity and emotional closeness can alter speech realization (Bänziger & Scherer, 2005; Costa et al., 2008; Farley et al., 2013), and this could have resulted in exaggerated pitch in speech produced in the interactions between these dyads. On the other hand, Knoll & Scharrer (2007) and Knoll et al. (2015) found a similar effect with higher mean pitch in FDS than NDS, which was not due to familiarity between interlocutors. Still, procedural differences may play a role. The 2007 study used a specific procedure, in which participants had to imagine non-native listeners and speak *as if* they were addressing them (see Section 5.1). In the 2015 study, all the participants had just been speaking to people with hearing loss. This could have elicited higher pitch and resulted in carry-over effects when they switched to using FDS.

In summary, this section reviewed evidence on vowel hyperarticulation, low speech rate, high intensity, and high pitch correlates in FDS (see Figure 2 for a summary of the FDS features). Most research has found vowel hyperarticulation and low speech rate in FDS. Conversely, there is less evidence of intensity and different pitch features (range, contour, and mean) being different between FDS and NDS. FDS mainly employs flat contours, but in some cases rising contours occur, resembling the contours of interrogative utterances. FDS may be characterized by a wide pitch range, which could reflect hyperarticulation, although little research effort has been dedicated to exploring this feature. Lastly, several studies failed to report higher mean pitch in FDS in comparisons to NDS, although some supportive evidence has been noted (e.g., Kendi & Khattab, 2019).

INSERT FIGURE 2 ABOUT HERE

3. Emotional valence of FDS

The emotional or affective properties of FDS have also received attention in the literature. In studies exploring this question, participants were asked to listen to continuous speech samples and rate how positive or negative they sounded. The emotional valence of a speech signal is a complex combination of several acoustic features including, but not limited to, speech tempo, pitch height and range, and intensity (see Liscombe et al., 2003; Tursunov et al., 2019). While the perceived affect of a speech register can be directly related to speakers' desire to transmit emotion or to their communicative intent, it can also be a by-product of the exaggeration of prosodic and acoustic components intended to enhance a register's clarity or its didactic purpose. As discussed below, these components can elicit negative perceptions from listeners, which in turn can have an effect on the register's effectiveness as a linguistic tool. Critically, perception of the emotional valence of FDS appears to vary depending on the linguistic profile of the listener, so next we separately consider studies in which ratings were provided by either foreign or native listeners.

Non-native listeners. Bobb et al. (2019) identified a positive correlation between median pitch in FDS and ratings of speakers' competence and friendliness (see Lynch, 1988). In this study, FDS consisted of sentences read to an imaginary audience rather than naturally produced speech to a foreigner. FDS was compared to other speech registers including NDS, clear speech, and IDS, yielding several interesting results. FDS was perceived as friendlier than NDS. IDS contained the highest level of median pitch, followed by FDS, clear speech, and NDS (IDS > FDS > clear > NDS), suggesting that positive emotional affect is (at least partially) driven by pitch height. Each register was produced by native speakers, but naïve (foreign) raters – who were not aware of speakers' language background and proficiency – considered speakers in the FDS condition to be overall less competent than speakers in the generic clear speech condition. On the other hand, speakers who produced higher median pitch in their FDS were rated as more competent and less condescending. Bobb et al. (2019) concluded that intelligibility did not positively correlate with perceived condescension, meaning that speaking clearly does not entail sounding condescending and patronizing.

Native listeners. Native listeners are expected to have different perceptions of FDS than the intended non-native audience because they would not derive any linguistic benefit from this register (unless they hear it in challenging listening conditions). That is, a positive or negative evaluation of FDS may depend on whether the listeners feel themselves to be the intended and appropriate addressees for the register adopted (Austerlitz, 1956; Ferguson, 1975). DePaulo and Coleman (1981) recruited 91 native English listeners to rate communications directed to non-native listeners (as well as to infants, adult native listeners, and people with cognitive disability). These participants perceived FDS as less friendly, less respectful, and less encouraging than NDS, but they considered speakers using FDS to be more competent than speakers of all the other speech registers. Surprisingly, in a later study using the same methodology and a combination of measures, DePaulo and Coleman (1987) instead found that FDS was considered to be warmer than NDS. The authors stressed that listeners displayed a remarkable ability to recognize differences between registers even without any explicit/external cues as to the identity of the addressees (see also Knoll et al., 2011a). This indicates that FDS is clearly differentiated from other registers, suggesting that it serves a communicative function and conveys psychological and sociological information (DePaulo and Coleman, 1981).

In more recent work, Uther et al. (2007), Knoll and Scharrer (2007), and Knoll et al. (2011b) used low-pass filtered segments of vocal interactions between two native speakers of English, a native speaker and their infant, a native speaker and a non-native confederate (a native Chinese speaker) to elicit ratings of negative and positive vocal affect from naïve native listeners. This band filter removed all frequencies above 1000Hz rendering speech unintelligible, while leaving prosodic and emotional features unaffected (Scherer, 1971; Scherer et al., 1972; Starkweather, 1967), so that raters had to rely on acoustic features for their emotional evaluations, without considering semantic content. In Uther et al. (2007) and Knoll and Scharrer (2007), FDS received the lowest ratings for positive vocal affect and the highest ratings for negative affect (see also Knoll et al., 2009a), whereas in Knoll et al. (2011b), FDS received lower rating than NDS for positive vocal affect only. In addition, Knoll et al (2009b) tested the consistency of emotional ratings across various band filters, in addition to the 1000 Hz cut-off, to measure the contribution of higher frequency bands to rating scores. They found that across most filter levels (except the unfiltered and the 1200Hz filter version), NDS was rated as having higher positive vocal affect than FDS, but it was considered to request the same level of attention.

A number of acoustic parameters may be responsible for the different impressions made by different registers, especially the negative perception of FDS by native listeners. For instance, it has been argued that speech rate (Stewart, Bouchard-Ryan, 1982; Knoll et al., 2009a), vowel hyperarticulation (Uther et al., 2007), and pitch (Knoll et al., 2015) modulate the degree of negativity associated with FDS (see Rothermich et al., 2019). Rothermich et al. (2019) reported that in Uther et al. (2007) more hyperarticulated vowels in IDS got more positive ratings, whereas in FDS greater vowel hyperarticulation corresponded to more negative ratings. High pitch and wide pitch range also seem to play a role in eliciting positive emotional evaluations (cf. Knoll et al., 2015). Knoll et al. (2009b) found that the positive vocal affect of IDS decreased as low-pass band filters excluded higher frequencies, therefore reducing the contribution from the high pitch cue. We know that high pitch is a typical feature of IDS, hence it is likely that the exclusion of higher frequencies was the cause of the reduction in the perceived positive vocal affect in this register. This indicates that high mean pitch of IDS has some influence over its positive affect. However, it is worth noting that in Knoll et al. (2009b) pitch was higher in NDS than FDS, and this could have accounted for its higher positive vocal affect ratings and partially account for the rating discrepancy between these two registers (Biersack et al., 2005; Uther et al., 2007).

Most experiments with native raters have found that FDS is perceived negatively, even when semantic content is obscured by using various band filters. However, studies to date have not assessed listeners' beliefs about the intended audience for each register that they were asked to rate. We suggest this assessment should be included in future perceptual studies: listeners' ratings may be influenced by their perceptions of the register used and its intended audience. For instance, when native raters think that FDS is addressed to them, they might find it condescending and rate it negatively; conversely, if they believed it was addressed to someone else (especially a non-native listener), they might rate it more positively. Since FDS is a register directed to adults, it could be used to convey a disrespectful message to native listeners and be negatively rated for this reason (see Starkweather, 1967; Clyne, 1981). It is likely – but this proposal needs deeper investigation – that low speech rate and vowel hyperarticulation in the absence of a positive emotional contribution from high mean pitch both play a role in eliciting negative judgments of FDS from native raters.

4. The theory behind FDS.

The discussion above indicates that speakers adjust their speech when addressing non-native listeners in a way that is proposed to assist speech processing and comprehension. Now, we turn to the theories that have tried to account for these accommodations in FDS. Earlier theoretical approaches hypothesized that FDS was an example of a simplified register, largely characterized by syntactic and lexical simplifications (Henzl, 1973; Tweissi, 1990). FDS was assumed to be the result of a communication strategy determined by cultural rules (see Ferguson, 1975; Canale & Swain, 1980; Tarone, 1980). More recently, FDS has generally been interpreted as the result of speech accommodation by L1 speakers who want to maintain successful communication with L2 listeners (Giles et al., 1991; Hazan et al., 2015; Scarborough et al., 2007; Smith, 2007; Snow et al., 1981; Zuengler, 1991; see also Costa et al., 2008). This view assumes that speakers are sensitive to the addressee's need to receive linguistic clarifications and learn phonological contrasts.

The Hyperarticulation & Hypoarticulation (H&H) theory of speech accommodation (Lindblom, 1990) is the main theoretical framework adopted to interpret FDS research findings in the recent literature. This theory sustains that the main source of speech variability is accommodation to listeners, situations, and contexts. According to this theory, speakers continuously regulate their speech production along the hypo-/hyper-articulation continuum, in order to meet their communicative aims, listeners' demands, and to maintain successful communication. The articulation continuum spans the range from least to most effortful articulation, where the least effort is put into interaction with peers (i.e., NDS), following the natural tendency to save articulatory energy as much as possible without losing category distinctiveness (effort-based approach to Optimality Theory, Kirchner, 1997; Theory of Adaptative Dispersion, Lindblom, 1990; Diehl & Lindblom, 2004). The Communication Accommodation Theory (CAT) offers an alternative but similar perspective that accounts for previous experience with interlocutors, non-linguistic cues (such as smiling), and adoption of listeners' communication behaviours (Beebe & Giles, 1984; Coupland et al., 1988; Giles, 2016; Giles et al., 1991; Zhang & Giles, 2017). The latter element is part of a so-called *convergence* strategy, which enhances communication success through modifications of segmental and suprasegmental properties. The opposite strategy is called *divergence*, and is designed to maintain social distance by eschewing speech adjustments and demonstrating indifference to effective communication (for additional theoretical frameworks, see Nekvapil & Sherman, 2015; Wooldridge, 2001; Zuengler, 1991).

Both the H&H and CAT speech accommodation frameworks entail that most FDS adjustments are regulated by didactic intentions. There is, therefore, widespread theoretical consensus on the didactic purpose of FDS (Biersack et al., 2005; Smith, 2007; Margić, 2017; Rothermich et al., 2019; Bobb et al., 2019). However, as we discuss in later sections of this review, despite this consensus, there is little direct evidence establishing that FDS is effective in achieving its proposed didactic impact, and whether any positive effects associated with this register actually enhance non-native listeners' subsequent L2 perception or production. In short, there is a pressing need for further research into the didactic intentions, functions, and impacts of FDS.

The next section of this review focuses on the factors that can influence or modulate acoustic adjustments and the proposed didactic purpose of FDS described in previous sections. According to these theoretical accounts, FDS speech accommodations are based on the listener, the speaker, the situation and communicative context. We discuss the effects of each these factors on the properties of FDS below.

5. Factors that influence speech accommodation.

5.1 Listener characteristics

Language proficiency and accentedness. Language proficiency and the accentedness of the listener may be the first factors that influence the level of accommodation in FDS and its acoustic realization. To our knowledge, no systematic investigations of the effect of listener's proficiency on FDS are available, but previous studies allow us to speculate that perceived proficiency modulates the extent to which native speakers are inclined to adjust their speech for L2 listeners. For instance, Snow and collaborators (1981) suggested that the perceived language proficiency of the non-native listener (in addition to their perceived social status and intelligence) is responsible for the magnitude of FDS effects (see also Gaies, 1979; Chaudron, 1978; Dahl, 1981; Liu, Kuhl, & Tsao, 2003 for evidence on IDS). Kühnert and Antolík (2017) results on different speech rate adaptations made by native French and English speakers (presented in Section 2.2) suggest accommodation differences across speakers largely depend on the listener's language proficiency. In fact, in this study, English listeners had about 3 years less L2 experience than their French counterparts (6.4 vs. 9.2 years, respectively). The authors acknowledged that this might be the reason why only French participants adapted their speech rate to help their listeners. Furthermore, other studies have found low speech rate in English FDS (Warren-Leubecker & Bohannon, 1982; Hazan et al., 2015), suggesting that Kühnert and Antolík's null result for English FDS was likely due to listener characteristics rather than a cross-linguistic difference in adaptation between English and French. In sum, speech rate may vary as a function of an addressee's proficiency and future research should aim to assess correlations between the foreigner listener's proficiency and the native speaker's speech rate in English and other languages.

Several studies suggest that accentedness has a tight negative correlation with language proficiency (Gallego, 1990; Kang et al., 2010; Munro & Derwing, 1995), meaning that a strong foreign accent corresponds to a low level of proficiency. While some low-proficiency addressees may have less obvious foreign accents (Munro & Derwing, 1995), it is still the case that speakers may be biased to interpret a strong L2 accent as a symptom of low proficiency, and consequently adapt their register to help these listeners (see Kang et al., 2010). To date, no experiments have explored proficiency versus accentedness in orthogonal designs; instead, experiments have relied on the generic perception of the confederate's strong foreign accent without differentiating between accent and proficiency (e.g., Uther et al., 2007). Furthermore, previous research has not provided many details or objective measurements of listeners' accentedness and/or proficiency with few exceptions (Lynch, 1988; Hazan et al., 2015; Kendi and Khattab, 2019). In the only study to specifically address accentedness, Lorge and Katsos (2019) asked a Greek confederate to emphasize her foreign accent while producing grammatically correct speech in English, yet even in this case, no measure of resulting accentedness was reported. In this study, speakers hyperarticulated vowels only when the confederate simulated a strong foreign accent. This suggests that accentedness has some influence on the properties of FDS.

These findings stress the importance of using consistent measurements across studies that allow for direct comparisons of proficiency and accentedness. If the listener's proficiency level affects the realization of FDS properties (i.e., speech rate) independently from accentedness, this would be strong evidence that this register has a generic didactic purpose. Conversely, if accentedness only drives FDS features, or some of them, FDS likely provides articulatory information to help highly proficient L2 listeners who nevertheless retain a strong foreign accent.

Perception of foreignness. Another listener characteristic that may influence FDS production is the perception of foreignness. This construct is largely linked to listeners' physical appearance, which can lead speakers to assume they are foreigners with linguistic difficulties even when there no evidence of strongly accented L2 speech or low L2 proficiency. Of course, it is noteworthy that a speaker's L1 and language background (e.g., bilingualism) are not directly related to their physical aspect, but some studies find that perception of foreignness is influenced by physical appearance, even when that is not justified by the interlocutor's language identity or proficiency (Bernstein et al., 2007; Ito et al., 2004; Rubin, 1992). If physical appearance (only) drives hyperarticulation in FDS (see Long, 1983), then the register would not serve a didactic purpose but rather reflect an intention to emphasise social distance and linguistic superiority (Biersack et al., 2005; Clyne, 1981; Valdman, 1981). Results from FDS ratings (Uther et al., 2007; Hazan et al., 2015) seem to be in line with this idea because native listeners might perceive FDS to be disrespectful and to have lower intelligibility than other clear registers (see Valdman, 1981 for a similar hypothesis). Ratings made by native listeners might be based on their perception of an imbalance in speaker-listener interactions; that is, adjustments that are not made to accommodate listeners, but rather due to prejudices about perceived foreignness.

To address this issue, Kangatharan et al. (2012) calculated the vowel hyperarticulation of a square vocalic area of speech directed to foreign-looking and native-looking confederates with or without foreign accents. They observed that the accent of the addressee did not induce vowel hyperarticulation, but the addressees' foreign appearance did. Surprisingly, native-looking listeners with foreign accents did not elicit any acoustic adjustments in speech. This finding, however, was not replicated by the same team in a study with different target vowels and a larger sample size (Kangatharan et al., 2015). Here, the physical aspect of the listeners had no effect on FDS, whereas a foreign accent elicited hyperarticulated vowels (Arthur et al., 1980). In order to shed light on this matter, future research on FDS is required to consider possible confounding factors derived from speakers' biases toward different ethnicities. At present, most studies conducted on FDS did not clearly state whether their confederates had both foreign physical appearances and foreign accents (i.e., Uther et al., 2007; Hazan et al., 2015; Kendi & Khattab, 2019), which hinders the distinction between the influence of these two factors. Based on this limited evidence, it appears that physical appearance alone, namely the perception of foreignness, is not sufficient to elicit FDS. However, further evidence that directly compares appearance and accentedness is required.

Imaginary addressees. Many studies have opted to use 'imaginary listeners' by eliciting FDS in the absence of a live non-native listener (Papaousek and Hwang, 1991; Biersarck et al., 2005; Scarborough et al. 2007; Knoll and Scharrer, 2007; Knoll et al., 2009a, 2009b, 2011a, 2015; Bobb et al., 2019). This approach is advantageous in terms of controlling interactions, by reducing the individual differences that inevitably arise from live face-to-face interactions. Still, this design choice prevents researchers from controlling for factors such as accentedness and language proficiency, which as discussed above, can influence the realization of FDS, and may explain some of the contradictory findings in the current literature. The validity of these paradigms may also be questioned: the interaction may not be as natural as one with a live interlocutor present and each participant might imagine a different 'stereotypical' foreigner, possibly depending on their previous personal experiences (Snow et al., 1981). Any of these imagined differences could elicit different degrees of acoustic adaptation in FDS.

In fact, findings from studies that employed imaginary addressees to elicit FDS have been mixed. Some studies have reported adjustments similar to those reported in the presence of real addressees (Biersack et al., 2005; Bobb et al., 2019; Scarborough et al., 2007), while others report no differences from NDS (Knoll et al., 2009a, 2011a; Knoll & Scharrer, 2007). Scarborough et al. (2007) compared the results of interactions with imaginary and real non-native listeners and found that the imaginary listeners elicited low speech rates to a greater degree than real addressees (see also Knoll et al., 2009a). This demonstrates that real and imaginary audiences can elicit unequal manifestations of FDS features (Knoll & Scharrer, 2007; Knoll et al., 2009a). It seems likely that imaginary addressees lead to inauthentic speech modifications, and that such adjustments likely vary across participants due to their own performance abilities (i.e., actresses vs. students in Knoll et al., 2009a), experience with L2 speakers, or the instructions they received on the experimental task (Snow, 1981; Lam et al., 2012; Knoll et al., 2011a). To counteract some of these issues, further research involving fictitious listeners could consider employing simulations of live interactions, i.e., making participants believe they are talking to a real foreigner. This option still does not require actual addressees, making it practically feasible, and in turn has several benefits. A simulation, for instance, guarantees stable comparisons across participants, thanks to the standardization of the fictitious listener's behaviour (see Buz et al., 2016). It also allows the researcher to control various factors such as the physical appearance and accentedness of the listener, simultaneously.

5.2 Speaker characteristics

Production of FDS acoustic features seems to be speaker dependent. In fact, Knoll et al. (2011a) compared students and actresses' FDS production, and observed that, with the same amount of exposure to non-native listeners, the actresses hyperarticulated vowels more than the students (see also Knoll et al., 2009a). In addition, the nature of the relationship between interlocutors might also induce speakers to tailor their speech to listeners' needs.

Previous experience and bilingualism. Experience communicating with L2 listeners is one factor that appears to increase speakers' sensitivity to listeners (Snow et al., 1981), and makes them more likely to accommodate their speech (i.e., to use FDS). For instance, language teachers, who are used to dealing with

L2 learners’ difficulties, might be particularly prone to employ effective speech adjustments, which would result in specific adaptations in their speech production matched to their students’ L1 phonological inventory. Another population that has been studied in this regard are bilingual speakers. Lorge & Katsos (2019) found that bilinguals hyperarticulated vowels more than monolinguals in FDS. This suggests that bilingualism shapes FDS, and that individuals who are immersed in multilingual environments, and who may have been L2 learners themselves, are more prone to use this speech register.

Emotional closeness, familiarity, and relationship. Speakers may differentially adjust their speech depending on the nature of their relationship with their interlocutor. People are likely to behave differently with elders as compared to same-age peers, with people whom they know, or have a close relationship to outside the experimental context such as romantic partners (Bänziger & Scherer, 2005; Caporael, 1981; Farley et al., 2013; Kemper et al., 1995). Young people or caregivers often overaccommodate their speech when talking to elders or people with disabilities, conveying condescension (Coupland et al., 1988; Ryan et al., 1994; Ryan et al., 1986). This demonstrates that both age and familiarity shape the relationship between interlocutors and influence their speech adjustments. Kendi and Khattab (2019) used foreign addressees (age not reported) who had been working for the participants of the study (the native speakers) for at least 2 months. Speakers and listeners knew each other, some for up to four years. Native speakers produced FDS characterized by wider vocalic space, limited to F1 movement, and higher pitch than NDS. This finding is in line with studies on articulation and pitch modulation in speech addressed to lovers or intimate friends. This suggests that social aspects of the relationship between interlocutors, like distance or closeness (i.e., relationships with superiors or peers) may be relevant to the way in which FDS is delivered. In line with this possibility, other work on FDS has employed strangers as listeners, and did not find higher mean pitch, but instead observed an expanded vocalic triangle manifested as both F1 and F2 exaggeration (Uther et al., 2007; Kangatharan et al., 2015). Note that, as far as we know, the latter set of studies used English as the target language, which may limit the generalization of these findings to other languages. Future studies should probe the influence of different types of relationships between interlocutors on FDS, especially in languages other than English, and consider age, social distance, emotional closeness, and immersion in a multilingual environment as factors that potentially play a role in shaping FDS features.

5.3 Situational and contextual factors

Situational factors: task instruction and communicative goal. Situational factors influence the properties of FDS and include, but are not limited to, the instructions given to the speakers to elicit FDS (Knoll et al., 2011a), and the purpose of the conversations in which they use FDS. There are numerous potentially important situational elements, and they are often tightly bundled together, making it difficult to disentangle their effects. At present there is still little evidence for the influence that experimental tasks have on the acoustic properties of FDS, but it is likely they have important effects. For instance, previous research indicates that the instructions used to elicit a clear register directly influence its realization (Lam et al., 2012; see also Smiljanić & Bradlow, 2009; Knoll et al., 2011a). Knoll et al. (2011a) investigated FDS features by employing “simulated free speech” and “standardized sentences”. In the former task, speakers were provided with three target toys (a shark, a sheep, and a shoe) so that they could invent their own scenarios (centred on the target words) to address imaginary listeners. In the latter task, speakers used fixed sentences like ”Look at the ‘target word’” to address the same imaginary listeners (e.g. “Look at the shark”). The authors observed differences in the acoustic features elicited in the two tasks and concluded that the reproduction of some FDS features depends on the task employed. Hence, the type of task may induce peculiar speech modifications and communicative strategies; for instance a task where participants have to read aloud (as in Bobb et al., 2019) likely results in FDS with peculiar phonetic characteristics that are different than those of spontaneous speech tasks (see Blaauw, 1992; Hazan & Baker, 2010; Laan, 1992). A *tandem* situation where two interlocutors practice language exchange (as in Kühnert & Antolík, 2017), provides for a free ranging, natural, conversational situation. By contrast, requiring a speaker to give directions to a listener over the phone (as in Smith, 2007) entails a strictly defined situation and a possibly hierarchical relationship between interlocutors. In the former case, the target audience probably feels freer to ask for repetitions when the speaker’s enunciation is not sufficiently clear. There are other factors, not directly related to listeners or speakers themselves, that are not fixed, but rather vary dynamically. For instance, it may be hypothesized that speakers provide slightly different acoustic cues depending on the immediate goals of communication; for instance, if they are using FDS to teach the proper pronunciation of phonemes or the spelling of ambiguous words. These different didactic goals may require different approaches to situational difficulties and possibly different articulatory strategies.

Contextual factors: feedback and perceived successful communication. Feedback from listeners during communication may induce speakers to dynamically regulate their speech rates or increase the perceptual distance between ambiguous phonemes in a specific word. In fact, feedback from the interlocutor seems to be fundamental in eliciting hyperarticulatory phenomena in communicative interactions (Buz et al., 2016; Maniwa et al., 2009; Ohala, 1984; Stent et al., 2008; see also Smith & Trainor, 2008 for evidence on IDS). Two studies demonstrate that speakers make dynamic adjustments to their speech properties in response to interlocutors' feedback (Burnham et al., 2010; Buz et al., 2016). Burnham et al. (2010) observed that speakers hyperarticulated vowels when addressing a computer avatar (rather than a human), who repeated their sentences, and they did so to a greater extent after the avatar pronounced their sentence incorrectly than when it did so correctly (see also Schertz, 2013 for a study on exaggerated contrasts after listener's misheard speech). Buz et al. (2016) observed similar results in a simulated native-native interaction where speakers hyperarticulated plosive consonants after receiving negative feedback, and then maintained this alteration across several trials. In fact, some speech adjustments emerge only if speakers perceive that successful communication is useful to the listener (Kuhlen & Brennan, 2013; Lockridge & Brennan, 2002). In short, feedback appears to shape speech registers because it induces the production of clear features such as vowel hyperarticulation.

With regards to FDS, only one study to date has assessed the role of feedback on its realization. Warren-Leubecker and Bohannon (1982) directly explored the role of feedback on the online adjustment of speech during native speaker and non-native listener interactions. They found that regardless of non-natives' L2 proficiency levels, lower FDS speech rates were mostly driven by feedback indicating that communication had failed. This result, together with previous findings from NDS studies, suggests that feedback may play a significant role in FDS production (see Suffill et al., 2021 for evidence on lexical alignment with non-native listeners). Alternatively, negative feedback could demonstrate comprehension difficulties, indicating the listener has low proficiency. In this case, FDS might be elicited mainly as a function of the listener's language proficiency level. To test these possibilities, future research should focus on disentangling main effects and the interaction between feedback and proficiency in order to define the role of each of those factors on FDS production. Warren-Leubecker and Bohannon's (1982) evidence on the feedback mechanism supports the H&H hypothesis (see Section 4) that FDS serves a didactic purpose in response to a listener's linguistic needs. However, future research should aim to establish whether didactic purpose and feedback determine the acoustic changes in FDS to a similar degree and whether their effects can indeed be disentangled to provide a more precise explanation for the observed properties of FDS.

In Section 5, we showed that FDS features are influenced by multiple factors, some related to listeners and others to speakers (see Figure 3 for a summary). FDS is the result of a complex set of factors, which include, for instance, adaptations to low proficiency listeners and the nature of the personal interaction between interlocutors. Lastly, we discussed that speech is adapted according to feedback from listeners in line with the goals of successful communication. The information presented in Section 5 indicates that FDS is not a static register, but rather adapts to situations, context, and interlocutors' interactions. The dynamic nature of FDS is strictly bound to its didactic purpose, and Section 6 provides a discussion of FDS features in relation to other clear speech registers that will further help to understand this aspect.

INSERT FIGURE 3 ABOUT HERE

6. Differences and commonalities between FDS and other clear speech styles

A comparison between FDS and IDS suggests that both registers serve a didactic function: they are produced to enhance language acquisition. By contrast, the clear features of Lombard Speech cannot serve this function; this register simply reflects the need to communicate clearly in a noisy environment. But similar acoustic adjustments might result from different underlying purposes, adaptation to listeners' needs, and communicative goals.

6.1 Comparing FDS and IDS

IDS is the speech register that adults use in interactions with young infants (Golinkoff et al., 2015). It has a number of linguistic, emotional, and acoustic characteristics that differentiate it from Adult Directed Speech (ADS, which is equivalent to NDS), including simplified grammar (Soderstrom, 2007), warm positive affect (Kitamura & Burnham, 2003), changes in speech timbre (Piazza et al., 2017), low speech rate (Panneton et al., 2006), exaggerated pitch height and range (Fernald et al., 1989), and acoustic exaggeration of vowels (Kuhl et al., 1997). Vowel hyperarticulation in IDS (Burnham et al., 2015; Cristia & Seidl, 2014; Kalashnikova & Burnham, 2018), has been proposed to serve a specific linguistic function, similar to FDS. Compared to ADS, caregivers using IDS significantly expand the acoustic space between the three corner vowels /i/, /u/, and /a/. This is proposed to result in clearer speech that helps infants discriminate the phonetic categories of their native language and later reproduce them in their own vocal tract. However, this proposal has been debated in IDS literature (Cristia, 2013). First, while vowel hyperarticulation in IDS has been reported for a number of languages including English (Adriaans & Swingley, 2017; Burnham et al., 2002; Kalashnikova & Burnham, 2018), Russian and Swedish (Kuhl, 1997), Spanish (García-Sierra et al., 2021), and Mandarin Chinese (Liu et al., 2009), it has not been detected in Dutch (Benders, 2013), German (Audibert & Falk, 2018), or Norwegian (Englund & Behne, 2005) IDS. Second, vowel categories in IDS are more variable than those in ADS, so despite the expansion of the space between corner vowels, overall vowel clarity is reduced, and non-corner vowel categories are less discriminable (Cristia & Seidl, 2014; Martin et al., 2015; McMurray et al., 2013). This evidence has led to the proposal that vowel hyperarticulation in IDS does not facilitate language acquisition but is instead a by-product of other affective adjustments made in this register such as changes in voice quality and smiling (Benders, 2013; Miyazawa et al., 2017).

Hence, it is possible that the acoustic exaggeration of vowels observed in IDS and FDS stem from different speaker intentions and articulatory mechanisms. Kalashnikova, Carignan, and Burnham (2017) provided the first direct evidence for this possibility. In their study, mothers spoke to an adult in a typical manner (ADS), to an adult in a clear and exaggerated manner (exaggerated speech, ES), and to their infant (IDS) while their tongue and lip movements were measured using electromagnetic articulography. Acoustic analyses of maternal speech indicated that ES and IDS contained more hyperarticulated vowels than ADS. However, mothers exaggerated their tongue movements, that is, actually hyperarticulated speech, only in ES and not IDS. Acoustic exaggeration of vowel F1 and F2 in IDS, was instead explained to result from the significantly greater reduction in the size of the vocal tract through laryngeal raising in IDS compared to both ES and ADS. This adjustment is typically observed when a speaker wants to appear smaller and less threatening. The authors proposed that the acoustic exaggeration of vowels in IDS may have originated as a by-product of a maternal intent to sound friendly and comfort infants. However, while not originally aimed at facilitating infants' language development, this 'accidental' component of IDS may have acquired a secondary linguistic function.

In line with this proposal, there is evidence that vowel exaggeration is modulated by infants' linguistic and processing needs, and that infants benefit from this component of speech input. First, reduced vowel exaggeration has been reported in IDS to infants who are unable to hear their mothers' speech (Lam & Kitamura, 2012), or when infants are at-risk for a language processing disorder such as dyslexia (Kalashnikova et al., 2018; 2020). Thus, mothers appear to adjust the vowel properties of their IDS to the specific needs of their infant audiences. Second, hyperarticulated vowel sounds elicit more mature neural responses and more successful sound discrimination in nine-month-old infants (Peter et al., 2016) and facilitate word recognition in 19-month-olds (Song et al., 2010). Critically, these relations are observed at the level of individual mother-infant dyads: mothers who exaggerate vowels to a greater extent in their IDS have infants with more advanced speech perception skills (Kalashnikova & Carreiras, 2021; Liu, Kuhl, & Tsao, 2003) as well as larger concurrent and future vocabularies (Hartman et al., 2017; Kalashnikova & Burnham, 2018; Lovcevic et al., 2020).

The prosodic components of IDS, namely slow rate, pitch height, and pitch range also facilitate speech processing and lead to positive language acquisition outcomes in young infants (Spinelli et al., 2017). This is interesting given that these speech components are typically associated with the affective function of the register and are not consistently found in other clear speech registers including FDS (Biersack et al., 2005; Uther et al., 2007; Knoll et al., 2009a, 2011a; 2011b; Lorge & Katsos, 2019; Bobb et al., 2019). For instance, speech stimuli with the prosodic properties of IDS have been shown to facilitate infants' neural encoding of speech (Kalashnikova, Peter, et al., 2018; Zangl & Mills, 2007), vowel discrimination (Trainor & Desjardins, 2002), segmentation of continuous speech (Thiessen et al., 2005), and word learning (Graf Estes & Hurley, 2013; Ma et al., 2011). Adults also benefit from these properties as they are more successful at learning novel words when they are produced in IDS than in ADS (Golinkoff & Alioto, 1995; Ma et al., 2020).

As can be seen, some but not all IDS components overlap with FDS, and these similarities and differences have been used to support the claim that these components can occur independently of each other and are dynamically adjusted according to the specific emotional and linguistic needs of each audience (Burnham et al., 2002; Kalashnikova, Goswami, et al., 2018; Uther et al., 2007). However, more recent research has identified more nuanced similarities and differences between these two registers that help us better understand their possible didactic functions and roles in facilitating language acquisition and processing. It appears that vowel hyperarticulation and low speech rate are manifested to a similar degree in FDS and IDS (Uther et al., 2007; Lorge & Katsos, 2019; Martin et al., 2016). The main difference between the registers consists of the lack of pitch exaggeration in FDS compared to IDS, particularly with regards to the exaggeration of pitch contours (Knoll et al., 2015) and overall pitch height (Uther et al., 2007).

This review suggests that IDS and FDS share several components that may assist speech processing and language learning in their intended audiences. Infants benefit from the acoustic components of IDS when they occur in isolation or in unison, and there is some evidence that these components can also lead to processing benefits in adults. However, the presence of individual components in IDS is modulated by infants' age and linguistic experience. It is plausible that similar effects due to language proficiency can be observed in FDS. In fact, neglecting the importance of L2 proficiency may have led to inconsistent findings regarding the individual components of FDS and how they facilitate L2 perception and learning.

6.2 Comparing FDS and LS

Lombard speech (LS) is a register elicited when speakers have to counter background noise (Lombard, 1911). Compared to NDS, its characteristic articulatory and acoustic features include loudness, articulatory effort, low speech rate, and hyperarticulation (Garnier et al., 2006; Garnier et al., 2018; Sankowska et al., 2011; Hazan et al., 2015). Most of these features are shared with FDS, including loudness, low speech rate, and hyperarticulation (Hazan et al., 2015; Sankowska et al., 2011), but research has also uncovered several key differences. With regards to loudness, some research reported FDS to be louder than NDS, as we noted in Section 2.3. However, to the best of our knowledge, there is one study that has compared intensity across LS and FDS and it found no significant difference between the registers (Hazan et al., 2015). Whereas the difference between FDS and NDS could be predicted, the lack of distinction between FDS and LS is surprising. In fact, we would instead expect LS to be significantly louder than FDS since the latter is not specifically intended to overcome noise. This predicted difference also aligns with accommodation theories, which predict speakers adjust to better accommodate listeners' needs.

Sankowska et al. (2011) explored other aspects of speech that distinguish FDS from LS. In this study, the authors found that FDS emphasizes phoneme duration contrasts that help distinguish short from long speech sounds, whereas LS emphasizes duration differences less than FDS. Hazan et al. (2015) compared NDS, FDS and two acoustic barriers, namely, vocoded and noisy speech. Compared to NDS, speakers modified their speech across all other conditions, but in the vocoded condition they lowered speech rate, lengthened words, and hyperarticulated vowels more than in FDS. As can be seen, LS and FDS share similar acoustic features, which are manifested to different extents. Specifically, the available results to date suggest that LS uses more hyperarticulated vowels and slower speech rates (lengthened words) than FDS, but FDS uses length contrastively to highlight phoneme differences to a greater extent than LS.

LS and NDS share a native audience, but LS is a clearer register produced to counteract interference that lowers the intelligibility of the message for native listeners, who otherwise (without interference) would understand it perfectly. In fact, LS is only designed to overcome acoustic interference; the addressee faces no linguistic difficulty and has no need to learn the language. In short, LS does not have a didactic purpose. This is in line with the differences described in the features of LS and FDS and the perception of LS by both native and non-native listeners. In fact, Cooke & Lecumberri (2012) discovered that non-native listeners are not able to take advantage of LS clarity like native speakers (see also Bradlow and Benet, 2002 for a similar effect on clear speech), suggesting that LS and FDS fulfil different functions.

The comparison between FDS and LS shows that the aims of communication and addressees are crucial for eliciting speech registers. The didactic function of FDS is not limited to hyperarticulation, given that LS also exhibits this feature, but it does not serve a didactic purpose. The fact that both FDS and LS are characterized by low speech rates and vowel hyperarticulation does not make them similar versions of clear speech. Rather, speech registers result from the sum of various factors such as the speaker's intention and the specific communicative goal (e.g., to overcome linguistic difficulties in the case of FDS vs. noise in the case of LS). These factors, together with the addressee's linguistic needs and identity, seem to be the most relevant factors in eliciting specific speech styles and their respective acoustic features (see Knoll et al., 2015 for similar results on FDS vs. speech directed to people with hearing impairments).

In Section 6, we discussed the differences and commonalities between FDS and two other clear registers. By comparing and contrasting acoustic features, we established that FDS and IDS are both likely to serve didactic purposes that nevertheless differ in some respects. In fact, we saw that the origin of the didactic function of these two registers is regulated by the specific needs of their audiences: addressee identity plays an important role in defining the characteristics and purpose of each register. As for FDS and LS, the evidence suggests that the two registers have highly similar acoustic features (loudness, low speech rate and vowel hyperarticulation), but that specific manifestations of these features may derive from different speaker intentions and listener needs. Loudness in LS is justified by its need to overcome background noise, which is not the case in FDS. Perceptual studies would help to untangle the differences and similarities between these registers in an objective manner. Ratings of clarity and other types of subjective ratings could help advance our understanding of the differences between these registers and their purposes (see Rothermich et al., 2019). With this in mind, the next section turns to existing research that has investigated the perception of FDS by native and non-native listeners.

7. Perception of FDS

Without the appropriate level of speech accommodation, non-native listeners experience frustration and lose interest in L2 learning (Zuengler, 1991; Kemper et al., 1995; Margić, 2017). The appreciation of FDS may depend on whether it meets their needs without being either overaccommodating or patronizing (Perdue, 1984; Coupland et al., 1988; Lindblom, 1990). According to the didactic view of FDS and our discussion above, both affect and clarity perceptions of FDS depend on the non-native listener's L2 proficiency (Chaudron, 1978; Dahl, 1981; Snow et al., 1981; Xu et al., 2013). However, studies on listener's perceptions of FDS are scarce and most focus only on the emotional valence of FDS using listener affective ratings (as discussed in Section 3). On the other hand, very few studies have focused on FDS intelligibility, or on how clear and useful L2 listeners consider FDS to be. Here we point out the most relevant results regarding the perceived clarity of FDS first by non-native listeners and then by native listeners.

Non-native listeners. Congruent with theories of accommodation and the didactic function hypothesis (Lindblom, 1990; Uther et al., 2007), non-native listeners tend to rate FDS as being clearer than NDS, possibly because FDS meets their needs for language learning (Hazan et al., 2015). In Bobb et al. (2019), participants had to assign clarity scores to FDS and NDS without knowledge of what register they were hearing; the non-native listeners rated FDS as clear speech, whereas NDS was rated as less intelligible. In Kangatharan et al. (2015), early and late L2 learners listened to samples of FDS and NDS with low to high levels of added noise and assessed their clarity by using a Likert scale. All participants perceived FDS to be clearer than NDS regardless of their L2 proficiency, and an interaction between noise level and speech register showed that FDS clarity was less affected by noise than NDS. Such results are crucial because they demonstrate that FDS is sharply differentiated from NDS (Depaulo & Coleman, 1981; Knoll et al., 2011a), and that it possibly boosts L2 intelligibility for non-native listeners. This is in line with the finding that non-native listeners do not consider LS to be as clear as native listeners do (Cooke & Lecumberri, 2012), supporting the assumption that LS lacks any didactic function. It appears that non-native listeners perceive FDS to be clearer than LS since only the former is intended to meet their linguistic and communicative needs. It is important to underline that, although most general features of FDS (e.g., low speech rate) likely enhance speech clarity for non-native listeners of any L1, it is also probable that this clarity effect partially depends on the non-native listeners' L1 and on whether production of FDS is oriented to accommodate listeners of that specific language.

The studies reported above indicate that FDS supports speech clarity at any level of L2 proficiency, but all conclusions are based on subjective survey ratings, and objective measurements of actual speech processing and comprehension are missing. Neuroimaging techniques would provide a way to obtain measurements that do not depend on raters' metacognitive skills and would directly assess the benefits of FDS. The only study to date that has employed this approach is Uther et al. (2012) who used electroencephalography (EEG). They recorded event related potentials (ERPs) derived from the perception of hyperarticulated words and measured mismatch negativity (MMN), which is an index of auditory discrimination. To assess whether vowel hyperarticulation helped L2 listeners to discern vowel contrasts, native and non-native listeners were tested in a word listening task, in which words were produced with either standard or hyperarticulated vowels. Results showed that the phonetic changes were detected regardless of the listener's language status: MMN was elicited by hyperarticulated vowels in both native and non-native listeners. This finding leaves open the questions of whether non-native listeners benefit from hyperarticulated vowels to perceive L2 phonemic contrasts and whether FDS enhances L2 perception as compared to NDS. In fact, the non-native participants had a high level of proficiency (about 9 years of L2 use) and were living in the country where their L2 was used at the time of the experiment. Hence, non-native participants had likely already acquired the phonological contrast used in the experiment, and they did not need hyperarticulation to aid its detection. Therefore, the question of whether L2 listeners benefit from vowel hyperarticulation remains open and requires further research, especially on non-native listeners with low levels of L2 proficiency.

Native listeners. Kangatharan et al. (2015) also asked native listeners to rate how clear they found FDS and NDS (with or without noise). Like non-native listeners, native listeners perceived FDS to be clearer than NDS, indicating that FDS is indeed a type of clear speech. In line with this finding, Hazan et al. (2015) explored native listeners' perception of clarity in FDS compared to NDS and Lombard Speech, which, as discussed above, exhibits a similar degree of vowel hyperarticulation as FDS but lacks its proposed didactic purpose. To do this, the authors used naturally elicited LS, FDS, and NDS (e.g., LS was elicited in a native-native conversation with added noise) and calculated the number of words produced by speakers to complete the task as an index of communicative difficulty. Hazan et al. (2015) reported that the speakers experienced greater communicative difficulty in the FDS compared to the LS condition (see Knoll et al., 2011a for similar results). Nevertheless, naïve raters who listened to those conversations in the absence of noise considered FDS to be clearer than NDS but less clear than LS. Hazan et al. (2015) offered the interesting explanation that register features that make speech clearer do not merely depend on the level of communicative difficulty. Such results support our discussion in Section 6.2 (comparing FDS and Lombard speech) by demonstrating that difficult listening conditions *per se* and the need for clarity are not sufficient for eliciting FDS.

Understudied aspects of FDS perception. In Hazan et al. (2015), native listeners and non-native listeners with low and mid-level proficiency rated FDS to be clearer than LS. Crucially, all three groups considered FDS to be clearer than NDS (as in Kangatharan et al., 2015), suggesting that FDS is perceived differently (and perceived to be clearer) at all levels of language proficiency. However, the lack of difference among the three listener groups, most importantly between the native and non-native listeners, does not offer support to the hypothesis that FDS has a didactic impact. In fact, research has not yet addressed whether FDS enhances language acquisition for L2 learners or whether, on the contrary, the sole way to improve L2 perception and production is exposure to native and peer to peer register (Margić, 2017). That is, no perceptual studies to date have explored the effects of FDS perception on L2 learning directly. As suggested above, research must address the effects of FDS exposure on non-native listeners' speech processing in order to understand its actual role in the process of L2 acquisition. If there is evidence that this register performs a didactic impact, non-native listeners would be expected to gain greater benefits from listening to FDS than native listeners, and to learn more when exposed to FDS compared to NDS. Note that perceptual ratings may also fail to highlight such differences because of intrinsic limits to subjective evaluations. One possibility is to expand research on FDS perception with neuroimaging techniques that provide more objective measurements of the phenomenon and to use them in combination with behavioural methods. We discuss this possibility in the next section.

8. Future directions and conclusion

This article reviewed the evidence for the FDS function of increasing speech intelligibility and facilitating L2 learning, by considering the main acoustic features and the factors influencing them. Low speech rate and vowel hyperarticulation were identified as the main features of FDS. We also examined research on additional acoustic features in FDS, such as wide pitch range and high intensity, which are still debated in the literature. Evidence revealed that FDS is a register based on listeners' identities, communicative needs and goals, and situational factors, such as the instructions provided for performing an experimental task. This suggests that FDS has a didactic purpose. Our discussion was grounded in the leading theoretical frameworks that account for the acoustic properties of FDS and supported by comparing FDS to two other clear speech registers, IDS and LS. We also reviewed empirical literature that has assessed the perception of FDS by native and non-native listeners, which yielded the following main findings. First, FDS is positively perceived only by non-native listeners, who also rate it to be clearer than NDS. Although native listeners rated FDS negatively, they still consider it to be clearer than NDS. This consensus further backs up the status of FDS as a clear speech register that is easily differentiated from NDS (Hazan et al., 2015; Uther et al., 2007). Second, FDS reduces vowel ambiguity in speech, which may provide listeners with useful information on how to perceive foreign sounds and (perhaps) produce them. Finally, clarity ratings of LS – a clear register meant to overcome communication noise – highlight that non-native listeners give LS lower clarity scores than native listeners. Taken together, this evidence from clarity ratings suggests FDS has a didactic impact in contrast to LS and NDS. However, further work is required to produce conclusive evidence for these possibilities and to understand whether non-native listeners benefit from FDS in the process of L2 acquisition. In this section, we discuss the open questions that we consider to be the most relevant directions for future research.

The first open question regards the typical acoustic characteristics of FDS (see Section 2), and whether these features are universally present in this register. For instance, it is not clear if speech rate is a feature of FDS present across languages, given that some research has shown mixed results for French and English (Smith, 2007; Kühnert & Antolík, 2017). Other features such as wide pitch range have simply not been adequately investigated in languages other than English to assess their universality. Further investigation is also needed on speaker status to determine whether speaker's identity (see Section 5) plays a role in eliciting FDS; it remains unclear whether all speakers produce FDS features to a comparable extent. Such factors could, for instance, include gender, based on evidence from one study (Lorge & Katsos, 2019) showing that women tend to hyperarticulate speech more than men. Another and highly interesting factor to investigate is speakers' bilingual status. It is plausible that bilinguals are better than monolinguals at adapting their speech to L2 listeners and at responding to audience needs and feedback (Lorge & Katsos, 2019). Moreover, it is probable that bilinguals who are also L2 teachers are particularly good at adapting their speech to their students' L1. Therefore, future experiments should investigate FDS production in bilinguals, expanding research in this field to languages other than English, and employing research designs that control for additional factors such as interlocutors' identities, listener's feedback, and adopt ecologically valid interactions for elicitation of FDS.

Relatedly, FDS features may be subject to other contextual variables such as the nature of the interaction, speaker familiarity, and communicative context. We hypothesize that contextual factors, such as listeners' feedback and communicative goals, are similarly relevant in determining the acoustic modifications in FDS, as described in Section 5. Evidence that communicative goals shape FDS, and that feedback due to miscomprehension induces further exaggeration of FDS features is essential for strengthening the claim that FDS serves a didactic purpose. To achieve this, future studies need to gather detailed demographic and linguistic information on both speakers and listeners, such as language background, proficiency, and accents.

Another outstanding question relates to the independence of the acoustic features of FDS. That is, one important venue for future research is to explore whether all the acoustic features of this register systematically co-occur, or whether they manifest independently from each other, serving different purposes for the speaker and the addressee. This is an important theoretical point for a better understanding of all clear registers, and FDS in particular. A parallel with other audience-oriented styles may point to an answer to this question (see Section 6). For instance, if we turn to developmental changes in IDS, we notice that several properties of IDS undergo drastic reshaping as the baby grows. In fact, pitch and speech rate in IDS seem to be adjusted to the infant's increasing age and linguistic ability, and become more adultlike in the second and third years of the child's life (Narayan & McDermott, 2016). Importantly, unlike other features, vowel hyperarticulation in IDS does not vary with infants' age, possibly reflecting the infants' continuing need to acquire language (Kalashnikova & Burnham, 2018; Liu et al., 2009). This suggests all the typical features of IDS do not need to co-occur and manifest to similar extents. This may also be the case for FDS. If future studies confirm the same pattern of vowel hyperarticulation and speech rate adjustments in FDS based on the characteristics of the listener (i.e., L2 proficiency and accentedness), this will reinforce the link between IDS and FDS and their didactic purposes.

Relatedly, the possibility that FDS may be characterized by a continuum of speech adjustments and accommodation requires attention in future research. Is FDS an on/off register, or does it occur on a continuum that goes from no adaptation (when the foreigner has a high level of phonological and linguistic competence) to the maximum grade of speech modifications with naïve L2 listeners. The results reported in Section 3 suggest that the latter may be the case, and that speech rate and vowel hyperarticulation might be modulated as a function of listeners' L2 proficiency. This aspect could be clarified via longitudinal studies on L2 acquisition and exposure to FDS, from naïve learners to proficient speakers. This may also reveal whether non-native listeners who are exposed to FDS benefit in terms of language learning (i.e., phonemes perception and pronunciation), as is proposed for IDS. If the continuum assumption is confirmed, this would constitute strong evidence for the didactic account of FDS and help confirm that FDS is the outcome of the speakers' unconscious goal to teach phonological contrasts to a non-native audience (Uther et al., 2007).

Combined behavioural and electrophysiological designs can offer an avenue for answering many open questions on FDS. The use of EEG can provide insight into the cognitive processes involved in listening to and interpreting hyperarticulated vowels and the effects of low speech rate on the brain activity of L1 and L2 speakers and listeners. As a complementary measure, ratings of speech segments by both native and non-native listeners would provide information on intelligibility, clarity, and emotional valence, which would be useful for interpreting the electrophysiological data. In addition, new techniques may assist the study of FDS production – which has never been assessed with neuroimaging techniques – because they efficiently limit the influence of muscular artifacts (Porcaro et al., 2015). Thus, we expect that future studies will make the most of available methods to investigate the cognitive processes of listener-oriented speech production as well. These future directions can be extended to include magnetoencephalography (MEG) and functional-Magnetic Resonance Imaging (fMRI), providing further fundamental information on the brain dynamics and localization of cognitive processes related to processing FDS.

To conclude, FDS likely boosts non-native listeners' speech comprehension. This is probably due to speakers' accommodating listeners' linguistic needs and results in adjustments such as vowel hyperarticulation and low speech rate. Nevertheless, further evidence that speakers adapt FDS to factors such as listener proficiency, listener feedback, and the specific aims of communication, is required to confirm theories that propose that FDS supports speech accommodation. Crucially, a deeper understanding of the factors that influence FDS production and perception is relevant for models of second language acquisition and can inform theoretical and practical approaches to second language instruction. If FDS serves a didactic purpose, it is imperative to assess how it benefits non-native listeners' perception and/or production of L2 phonological contrasts, and more generally, helps them learn their L2. Establishing that naïve L2 listeners appreciate and learn better when exposed to FDS would suggest that FDS is an important tool for second language teaching and for understanding all listener-oriented registers.

Supplementary material

A supplementary table summarizing designs and findings of the published studies that have assessed vowel hyperarticulation, speech rate, or pitch correlates of FDS can be consulted at https://osf.io/ndhr2/?view_only=baf8920dde914076b854ff322b499959

Acknowledgements

The authors would like to thank Magda Altman and Jessi Jacobsen for their help in editing the manuscript.

References

Adriaans, F., & Swingley, D. (2017). Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. *The Journal of the Acoustical Society of America*, *141*(5), 3070–3078. https://doi.org/10.1121/1.4982246

Al-Tamami, J., & Ferragne, E. (2005). Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. *INTERSPEECH-2005*, 2465–2468.

Andruski, J. E., Kuhl, P. K., & Hayashi, A. (1999). Point vowels in Japanese mothers' speech to infants and adults. *The Journal of the Acoustical Society of America*, *105*(2), 1095–1096. https://doi.org/10.1121/1.425135

Arthur, B., Wemer, R., Culver, M., Lee, Y. J., & Thomas, D. (1980). The Register of Impersonal Discourse to Foreigners: Verbal Adjustments to a Foreign Accent. *Discourse Analysis in Second Language Research*, 111–124.

Audibert, N., & Falk, S. (2018). Vowel space and f0 characteristics of infant-directed singing and speech. *9th International Conference on Speech Prosody 2018*, 153–157. https://doi.org/10.21437/SpeechProsody.2018-31

Austerlitz, R. (1956). Gilyak Nursery Words. *WORD*, *12*(2), 260–279. https://doi.org/10.1080/00437956.1956.11659604

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, *46*(3–4), 252–267. https://doi.org/10.1016/j.specom.2005.02.016

Beebe, L. M., & Giles, H. (1984). Speech-accommodation theories: A discussion in terms of second-language acquisition. *International Journal of the Sociology of Language*, *46*, 5–32. https://doi.org/10.1515/ijsl.1984.46.5

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behavior & Development*, *36*(4), 847–862. https://doi.org/10.1016/j.infbeh.2013.09.001

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, *18*(8), 706–712. https://doi.org/10.1111/j.1467-9280.2007.01964.x

Biersack, S., Kempe, V., & Knapton, L. (2005). Fine-tuning speech registers: A comparison of the prosodic features of child-directed and foreigner-directed speech. *INTERSPEECH 2005*. https://doi.org/10.13140/2.1.2133.5049

Blaauw, E. (1992). Phonetic differences between read and spontaneous speech. *ICSLP-1992*, 751–754.

Bloomfield, L. (1933). *Language*. Holt, Rinehart &Winston.

Bobb, S. C., Mello, K., Turco, E., Lemes, L., Fernandez, E., & Rothermich, K. (2019). Second Language Learners' Listener Impressions of Foreigner-Directed Speech. *Journal of Speech, Language, and Hearing Research: JSLHR*, *62*(9), 3135–3148. https://doi.org/10.1044/2019_JSLHR-S-18-0392

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *J. Acoust. Soc. Am.*, *112*(1), 13.

Bradlow, A. R., Kraus, N., & Hayes, E. (2003). Speaking clearly for children with learning disabilities: Sentence perception in noise. *Journal of Speech, Language, and Hearing Research: JSLHR*, *46*(1), 80–97. https://doi.org/10.1044/1092-4388(2003/007)

Burnham, D., Joeffry, S., & Rice, L. (2010). "D-o-e-s-Not-C-o-m-p-u-t-e": Vowel Hyperarticulation in Speech to an Auditory-Visual Avatar. *Anthrozoos A Multidisciplinary Journal of The Interactions of People & Animals*, *26*(3), 373–380. https://doi.org/10.2752/175303713x13697429463592

Burnham, D., Kim, J., Davis, C., Ciocca, V., Schoknecht, C., Kasisopa, B., & Luksaneeyanawin, S. (2011). Are tones phones? *Journal of Experimental Child Psychology*, *108*(4), 693–712. https://doi.org/10.1016/j.jecp.2010.07.008

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's New, Pussycat? On Talking to Babies and Animals. *Science*, *296*(5572), 1435–1435. https://doi.org/10.1126/science.1069587

Burnham, E. B., Wieland, E. A., Kondaurova, M. V., McAuley, J. D., Bergeson, T. R., & Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 years of age. *Journal of Speech, Language, and Hearing Research: JSLHR*, *58*(2), 241–253. https://doi.org/10.1044/2015_JSLHR-S-13-0205

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations. *Journal of Memory and Language*, *89*, 68–86. https://doi.org/10.1016/j.jml.2015.12.009

Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, *I*. https://doi.org/10.1093/applin/I.1.1

Caporael, L. R. (1981). The paralanguage of caregiving: Baby talk to the institutionalized aged. *Journal of Personality and Social Psychology*, *40*(5), 876–884. https://doi.org/10.1037/0022-3514.40.5.876

Chaudron, C. (1978). Foreigner talk in the classroom—An aid to learning. *Classroom Oriented Research in Second Language Acquisition*, 127–143.

Chaudron, C. (1979). Complexity of teacher speech and vocabulary explanation/elaboration. *In 13th Annual TESOL Convention*.

Clyne, M. G. (1981). 'Second Generation' Foreigner Talk in Australia. *International Journal of the Sociology of Language*, *1981*(28), 69–80. https://doi.org/10.1515/ijsl.1981.28.69

Cooke, M., & Lecumberri, M. L. G. (2012). The intelligibility of Lombard speech for non-native listeners. *The Journal of the Acoustical Society of America*, *132*(2), 1120–1129. https://doi.org/10.1121/1.4732062

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, *23*(4), 528–556. https://doi.org/10.1080/01690960801920545

Coupland, N., Coupland, J., Giles, H., & Henwood, K. (1988). Accommodating the elderly: Invoking and extending a theory. *Language in Society*, *17*(1), 1–41. https://doi.org/10.1017/S0047404500012574

Cristia, A. (2013). Input to Language: The Phonetics and Perception of Infant-Directed Speech: The Phonetics and Perception of Infant-Directed Speech. *Language and Linguistics Compass*, *7*(3), 157–170. https://doi.org/10.1111/lnc3.12015

Cristia, A., & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, *41*(4), 913–934. https://doi.org/10.1017/S0305000912000669

Dahl, D. A. (1981). The role of experience in speech modifications for second language learners. *Minn. Pap. Linguist. Philos. Lang.*, *7*(2), 78–93.

Depaulo, B. M., & Coleman, L. M. (1981). Evidence for the Specialness of the “Baby Talk” Register. *Language and Speech*, *24*(3), 223–231. https://doi.org/10.1177/002383098102400302

DePaulo, B. M., & Coleman, L. M. (1987). Verbal and nonverbal communication of warmth to children, foreigners, and retarded adults. *Journal of Nonverbal Behavior*, *11*(2), 75–88. https://doi.org/10.1007/BF00990959

Diehl, R. L., & Lindblom, B. (2004). Explaining the Structure of Feature and Phoneme Inventories: The Role of Auditory Distinctiveness. In S. Greenberg, W. A. Ainsworth, A. N. Popper, & R. R. Fay (Eds.), *Speech Processing in the Auditory System* (pp. 101–162). Springer. https://doi.org/10.1007/0-387-21575-1_3

Dodane, C., & Al-Tamimi, J. (2007). An acoustic comparison of vowel systems in adult-directed-speech and child-directed speech. *16th International Congress of Phonetics Sciences (ICPhS), Saarbrücken, 6-10 Août 2007*. https://hal.archives-ouvertes.fr/hal-03068849

Englund, K. T., & Behne, D. M. (2005). Infant Directed Speech in Natural Interaction—Norwegian Vowel Quantity and Quality. *Journal of Psycholinguistic Research*, *34*(3), 259–280. https://doi.org/10.1007/s10936-005-3640-7

Farley, S. D., Hughes, S. M., & LaFayette, J. N. (2013). People Will Know We Are in Love: Evidence of Differences Between Vocal Samples Directed Toward Lovers and Friends. *Journal of Nonverbal Behavior*, *37*(3), 123–138. https://doi.org/10.1007/s10919-013-0151-3

Ferguson, C. A. (1971). Absence of Copula and the Notion of Simplicity. A Study of Normal Speech, Baby Talk, Foreigner Talk and Pidgins. *Cambridge University Press*, 141–150.

Ferguson, C. A. (1975). Toward a Characterization of English Foreigner Talk. *Anthropological Linguistics*, *17*(1), 1–14.

Ferguson, C. A. (1981). 'Foreigner Talk' as the Name of a Simplified Register. *International Journal of the Sociology of Language*, *1981*(28), 9–18. https://doi.org/10.1515/ijsl.1981.28.9

Ferguson, H. S., & Quené, H. (2014). Acoustic correlates of vowel intelligibility in clear and conversational speech for young normal-hearing and elderly hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *135*(6), 3570–3584. https://doi.org/10.1121/1.4874596

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *J. Acoust. Soc. Am.*, *112*(1), 13.

Fernald, A. (1989). Intonation and Communicative Intent in Mothers' Speech to Infants: Is the Melody the Message? *Child Development*, *60*(6), 1497–1510. https://doi.org/10.2307/1130938

Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development*, *10*(3), 279–293. https://doi.org/10.1016/0163-6383(87)90017-8

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501. https://doi.org/10.1017/s0305000900010679

Gaies, S. J. (1979). Linguistic Imput in Formal Second Language Learning: The Issues of Syntactic Gradation and Readability in ESL Materials. *TESOL Quarterly*, *13*(1), 41–50. https://doi.org/10.2307/3585974

Gallego, J. C. (1990). The Intelligibility of Three Nonnative English-Speaking Teaching Assistants: An Analysis of Student-Reported Communication Breakdowns. *Issues in Applied Linguistics*, *1*(2), 219–250.

García-Sierra, A., Ramírez-Esparza, N., Wig, N., & Robertson, D. (2021). Language learning as a function of infant directed speech (IDS) in Spanish: Testing neural commitment using the positive-MMR. *Brain and Language*, *212*, 104890. https://doi.org/10.1016/j.bandl.2020.104890

Garnier, M., Bailly, L., Dohen, M., Welby, P., & Loevenbruck, H. (2006). An acoustic and articulatory study of Lombard speech: Global effects on the utterance. *INTERSPEECH*.

Garnier, M., Ménard, L., & Alexandre, B. (2018). Hyper-articulation in Lombard speech: An active communicative strategy to enhance visible speech cues? *The Journal of the Acoustical Society of America*, *144*(2), 1059–1074. https://doi.org/10.1121/1.5051321

Giles, H. (2016). Communication Accommodation Theory. In *The International Encyclopedia of Communication Theory and Philosophy* (pp. 1–7). American Cancer Society. https://doi.org/10.1002/9781118766804.wbiect056

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). Editions de la Maison des Sciences de l’Homme. https://doi.org/10.1017/CBO9780511663673.001

Golinkoff, R. M., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition*. *Journal of Child Language*, *22*(3), 703–726. https://doi.org/10.1017/S0305000900010011

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Graf Estes, K., & Hurley, K. (2013). Infant-Directed Prosody Helps Infants Map Sounds to Meanings. *Infancy*, *18*(5), 797–824. https://doi.org/10.1111/infa.12006

Han, M., de Jong, N. H., & Kager, R. (2018). Lexical Tones in Mandarin Chinese Infant-Directed Speech: Age-Related Changes in the Second Year of Life. *Frontiers in Psychology*, *9*, 434. https://doi.org/10.3389/fpsyg.2018.00434

Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language*, *44*(5), 1140–1162. https://doi.org/10.1017/S0305000916000520

Hatch, E. (1979). Interaction, input and communication strategies. *First Nordic Symposium on Interlanguage*.

Hatch, E., Shapira, R., & Gough, J. (1978). “Foreigner-Talk” Discourse. *ITL - International Journal of Applied Linguistics*, *39*(1), 39–60. https://doi.org/10.1075/itl.39-40.03hat

Hazan, V., & Baker, R. (2010). Does reading clearly produce the same acoustic-phonetic modifications as spontaneous speech in a clear speaking style? *Proceedings of DiSS-LPSS Joint Workshop 2010*, 7–10.

Hazan, V., Uther, M., & Granlund, S. (2015). How does foreigner-directed speech differ from other forms of listener-directed clear speaking styles? *ICPhS 2015*.

Henzl, V. M. (1973). Linguistic Register of Foreign Language Instruction. *Language Learning*, *23*(2), 207–222. https://doi.org/10.1111/j.1467-1770.1973.tb00656.x

Ito, T. A., Thompson, E., & Cacioppo, J. T. (2004). Tracking the Timecourse of Social Perception: The Effects of Racial Cues on Event-Related Brain Potentials. *Personality and Social Psychology Bulletin*, *30*(10), 1267–1280. https://doi.org/10.1177/0146167204264335

Kalashnikova, M., & Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language*, *45*(5), 1035–1053. https://doi.org/10.1017/S0305000917000629

Kalashnikova, M., Carignan, C., & Burnham, D. (2017). The origins of babytalk: Smiling, teaching or social convergence? *Royal Society Open Science*, *4*(8), 170306. https://doi.org/10.1098/rsos.170306

Kalashnikova, M., & Carreiras, M. (2021). Input quality and speech perception development in bilingual infants’ first year of life. *Child Development*. https://doi.org/10.1111/cdev.13686

Kalashnikova, M., Goswami, U., & Burnham, D. (2018). Mothers speak differently to infants at-risk for dyslexia. *Developmental Science*, *21*(1). https://doi.org/10.1111/desc.12487

Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports*, *8*(1), 13745. https://doi.org/10.1038/s41598-018-32150-6

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal*, *94*(4), 554–566. https://doi.org/10.1111/j.1540-4781.2010.01091.x

Kangatharan, J. (2015). *The role of vowel hyperarticulation in clear speech to foreigners and infants* [Brunel University London]. https://bura.brunel.ac.uk/bitstream/2438/10519/1/FulltextThesis.pdf

Kangatharan, J., Uther, M., Kuhn, L., & Gobet, F. (2012). A-M I S-P-EA-K-I-NG C- L- E- AR- L- Y E-N-OU-GH?: An investigation of the possible role of vowel hyperarticulation in speech communication. In S. F. d'Acoustique (Ed.), *Acoustics 2012*. https://hal.archives-ouvertes.fr/hal-00811042

Kemper, S., Vandepute, D., Rice, K., Cheung, H., & Gubarchuk, J. (1995). Speech Adjustments to Aging During a Referential Communication Task. *Journal of Language and Social Psychology*, *14*(1–2), 40–59. https://doi.org/10.1177/0261927X95141003

Kendi, A. A., & Khattab, G. (2019). Acoustic Properties of Foreigner Directed Speech. *International Congress of Phonetic Sciences 2019*, 5.

Kirchner, R. M. (1997). An effort-based approach to consonant lenition. *Doctoral Dissertation, University of California*.

Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy*, *4*(1), 85–110. https://doi.org/10.1207/S15327078IN0401_5

Knoll, M. A., & Costall, A. (2015). Characterising F(0) contour shape in infant- and foreigner-directed speech. *Speech Communication*, *66*, 231–243. https://doi.org/10.1016/j.specom.2014.10.007

Knoll, M. A., Johnstone, M., & Blakely, C. (2015). Can you hear me? Acoustic modifications in speech directed to foreigners and hearing-impaired people. *INTERSPEECH 2015*, 2987–2990.

Knoll, M. A., & Scharrer, L. (2007). Acoustic and Affective Comparisons of Natural and Imaginary Infant-, Foreigner- and Adult-Directed Speech. *INTERSPEECH 2007*, 4.

Knoll, M. A., Scharrer, L., & Costall, A. (2009a). Are actresses better simulators than female students? The effects of simulation on prosodic modifications of infant- and foreigner-directed speech. *Speech Communication*, *51*(3), 296–305. https://doi.org/10.1016/j.specom.2008.10.001

Knoll, M. A., Scharrer, L., & Costall, A. (2011a). “Look at the shark”: Evaluation of student- and actress-produced standardised sentences of infant- and foreigner-directed speech. *Speech Communication*, *53*(1), 12–22. https://doi.org/10.1016/j.specom.2010.08.004

Knoll, M. A., Uther, M., & Costall, A. (2009b). Effects of low-pass filtering on the judgment of vocal affect in speech directed to infants, adults and foreigners. *Speech Communication*, *51*(3), 210–216. https://doi.org/10.1016/j.specom.2008.08.001

Knoll, M. A., Uther, M., & Costall, A. (2011b). Using the Internet for speech research: An evaluative study examining affect in speech. *Behaviour & Information Technology*, *30*(6), 845–851. https://doi.org/10.1080/0144929X.2011.577192

Knoll, M. A., Walsh, S. A., MacLeod, N., O’Neill, M., & Uther, M. (2007). *Good performers know their audience! Identification and characterisation of pitch contours in infant and foreigner-directed speech.: Vol. Automated taxon recognition in systematics: Theory, approaches and applications,* (CRC Press/The Systematics Association, Boca Raton). MacLeod, N. (Ed.).

Knoll, M., Uther, M., MacLeod, N., O’Neill, M., & Walsh, S. (2006). Emotional, linguistic or just cute? The function of pitch contours in infant- and foreigner-directed speech. *Speech Prosody 2006*, 5.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America*, *115*(1), 362–378. https://doi.org/10.1121/1.1635842

Kuhl, P. K. (1997). Cross-Language Analysis of Phonetic Units in Language Addressed to Infants. *Science*, *277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, *20*(1), 54–72. https://doi.org/10.3758/s13423-012-0341-8

Kühnert, B., & Kocjančič Antolík, T. (2017). Patterns of articulation rate variation in English/French tandem interactions. In J. Volin & R. Skarnitzl (Eds.), *Pronunciation of English by Speakers of Other Languages*. Cambridge Scholars Publishing. https://halshs.archives-ouvertes.fr/halshs-01505928

Laan, G. P. M. (1992). Perceptual differences between spontaneous and read aloud speech. *Proceedings of the Institute of Phonetic Sciences Amsterdam*, 16.

Ladefoged, P. (2006). *A Course in Phonetics*. Thomson, Wadsworth. https://books.google.es/books/about/A_Course_in_Phonetics.html?id=jRmoQgAACAAJ&redir_esc=y

Ladefoged, P., & Maddieson, I. (1990). Vowels of the world's languages. *Journal of Phonetics*, *18*(2), 93–122. https://doi.org/10.1016/S0095-4470(19)30396-1

Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation: Hearing loss and vowel hyperarticulation. *Developmental Science*, *15*(2), 212–221. https://doi.org/10.1111/j.1467-7687.2011.01118.x

Lam, J., Tjaden, K., & Wilding, G. (2012). Acoustics of Clear Speech: Effect of Instruction. *Journal of Speech, Language, and Hearing Research*, *55*(6), 1807–1821. https://doi.org/10.1044/1092-4388(2012/11-0154)

Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 403–439). Springer Netherlands. https://doi.org/10.1007/978-94-009-2037-8_16

Liscombe, J., Venditti, J. J., & Hirschberg, J. (2003). Classifying subject ratings of emotional speech using acoustic features. *INTERSPEECH 2003*.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1–F10. https://doi.org/10.1111/1467-7687.00275

Liu, H.-M., Tsao, Feng-Ming, & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Developmental Psychology*, *43*(4), 912. https://doi.org/10.1037/0012-1649.43.4.912

Liu, H.-M., Tsao, F.-M., & Kuhl, P. K. (2009). Age-related Changes in Acoustic Modifications of Mandarin Maternal Speech to Preverbal Infants and Five-Year-Old Children: A Longitudinal Study. *Journal of Child Language*, *36*(4), 909. https://doi.org/10.1017/S030500090800929X

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, *9*(3), 550–557. https://doi.org/10.3758/BF03196312

Lombard, E. (1911). Le signe de l'élévation de la voix. *Annales Des Maladies de L'Oreille et Du Larynx.*, *37*, 101–119.

Long, M. H. (1981). Questions in foreigner talk disclosure. *Language Learning*, *31*(1), 135–157. https://doi.org/10.1111/j.1467-1770.1981.tb01376.x

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*(2), 126–141. https://doi.org/10.1093/applin/4.2.126

Lorge, I., & Katsos, N. (2019). Listener-adapted speech: Bilinguals adapt in a more sensitive way. *Linguistic Approaches to Bilingualism*, *9*(3), 376–397. https://doi.org/10.1075/lab.16054.lor

Lovcevic, I., Kalashnikova, M., & Burnham, D. (2020). Acoustic features of infant-directed speech to infants with hearing loss. *The Journal of the Acoustical Society of America*, *148*(6), 3399–3416. https://doi.org/10.1121/10.0002641

Lynch, A. J. (1988). Speaking up or talking down: Foreign learners' reactions to teacher talk. *ELT Journal*, *42*(2), 109–116. https://doi.org/10.1093/elt/42.2.109

Ma, W., Golinkoff, R. M., Houston, D. M., & Hirsh-Pasek, K. (2011). Word Learning in Infant- and Adult-Directed Speech. *Language Learning and Development*, *7*(3), 185–201. https://doi.org/10.1080/15475441.2011.579839

Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America*, *125*(6), 3962–3973. https://doi.org/10.1121/1.2990715

Margić, B. (2017). Communication courtesy or condescension? Linguistic accommodation of native to non-native speakers of English. *Journal of English as a Lingua Franca*, *6*. https://doi.org/10.1515/jelf-2017-0006

Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers Speak Less Clearly to Infants Than to Adults: A Comprehensive Test of the Hyperarticulation Hypothesis. *Psychological Science*, *26*(3), 341–347. https://doi.org/10.1177/0956797614562453

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., & McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition*, *129*(2). https://doi.org/10.1016/j.cognition.2013.07.015

Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition*, *166*, 84–93. https://doi.org/10.1016/j.cognition.2017.05.003

Munro, M. J., & Derwing, T. M. (1995). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, *45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *The Journal of the Acoustical Society of America*, *139*(3), 1272–1281. https://doi.org/10.1121/1.4944634

Nekvapil, J., & Sherman, T. (2015). An introduction: Language Management Theory in Language Policy and Planning. *International Journal of the Sociology of Language*, *2015*(232), 1–12. https://doi.org/10.1515/ijsl-2014-0039

Nelson, D. K. (1992). *The Foreigner Talk of a Family Physician: An Observational Study*. https://eric.ed.gov/?id=ED353826

Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica*, *41*(1), 1–16. https://doi.org/10.1159/000261706

Panneton, R., Kitamura, C., Mattock, K., & Burnham, D. (2006). Slow Speech Enhances Younger But Not Older Infants' Perception of Vocal Emotion. *Research in Human Development*, *3*(1), 7–19. https://doi.org/10.1207/s15427617rhd0301_2

Papoušek, M., & Hwang, S.-F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics*, *12*(4), 481–504. https://doi.org/10.1017/S0142716400005889

Peter, V., Kalashnikova, M., Santos, A., & Burnham, D. (2016). Mature neural responses to Infant-Directed Speech but not Adult-Directed Speech in Pre-Verbal Infants. *Scientific Reports*, *6*(1), 34273. https://doi.org/10.1038/srep34273

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers Consistently Alter Their Unique Vocal Fingerprints When Communicating with Infants. *Current Biology: CB*, *27*(20), 3162-3167.e3. https://doi.org/10.1016/j.cub.2017.08.074

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, *29*(4), 434–446. https://doi.org/10.1044/jshr.2904.434

Porcaro, C., Medaglia, M. T., & Krott, A. (2015). Removing speech artifacts from electroencephalographic recordings during overt picture naming. *NeuroImage*, *105*, 171–180. https://doi.org/10.1016/j.neuroimage.2014.10.049

Ramamurti, R. (1980). Strategies Involved in Talking to a Foreigner. *PENN Review of Linguistics*, *4*, 84–93.

Rodriguez-Cuadrado, S., Baus, C., & Costa, A. (2018). Foreign talk through word reduction in native/non-native spoken interaction. *Bilingualism: Language and Cognition*, *21*(2), 419–426. https://doi.org/doi:10.1017/S1366728917000402

Rothermich, K., Harris, H. L., Sewell, K., & Bobb, S. C. (2019). Listener impressions of foreigner-directed speech: A systematic review. *Speech Communication*, *112*, 22–29. https://doi.org/10.1016/j.specom.2019.07.002

Rubin, D. L. (1992). Nonlanguage Factors Affecting Undergraduates' Judgments of Nonnative English-Speaking Teaching Assistants. *Research in Higher Education*, *33*(4), 511–531.

Ryan, E. B., Giles, H., Bartolucci, G., & Henwood, K. (1986). Psycholinguistic and social psychological components of communication by and with the elderly. *Language & Communication*, *6*(1–2), 1–24. https://doi.org/10.1016/0271-5309(86)90002-9

Ryan, E. B., Hamilton, J. M., & See, S. K. (1994). Patronizing the old: How do younger and older adults respond to baby talk in the nursing home? *International Journal of Aging & Human Development*, *39*(1), 21–32. https://doi.org/10.2190/M52C-M2D2-R6C2-3PBM

Saadah, E. (2011). The production of Arabic vowels by English L2 learners and heritage speakers of Arabic. *Doctoral Dissertation, University of Illinois at Urbana-Champaign*.

Sankowska, J., García Lecumberri, M. L., & Cooke, M. (2011). Interaction of intrinsic vowel and consonant durational correlates with foreigner directed speech. *Poznań Studies in Contemporary Linguistics*, *47*. https://doi.org/10.2478/psicl-2011-0009

Scarborough, R., Dmitrieva, O., Hall-Lew, L., Zhao, Y., & Brenier, J. (2007). An acoustic study of real and imagined foreigner-directed speech. *The Journal of the Acoustical Society of America*, *121*(5), 3044–3044. https://doi.org/10.1121/1.4781735

Scherer, K. R. (1971). Randomized splicing: A note on a simple technique for masking speech content. *Journal of Experimental Research in Personality*, *5*(2), 155–159.

Scherer, K. R., Koivumaki, J., & Rosenthal, R. (1972). Minimal cues in the vocal communication of affect: Judging emotions from content-masked speech. *Journal of Psycholinguistic Research*, *1*(3), 269–285. https://doi.org/10.1007/BF01074443

Schertz, J. (2013). Exaggeration of featural contrasts in clarifications of misheard speech in English. *Journal of Phonetics*, *41*(3), 249–263. https://doi.org/10.1016/j.wocn.2013.03.007

Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, *118*(3), 1677–1688. https://doi.org/10.1121/1.2000788

Smiljanić, R., & Bradlow, A. R. (2009). Speaking and Hearing Clearly: Talker and Listener Factors in Speaking Style Changes. *Language and Linguistics Compass*, *3*(1), 236–264. https://doi.org/10.1111/j.1749-818X.2008.00112.x

Smith, C. (2007). Prosodic accomodation by French speakers to non-native interlocutor. *ICPhS XVI*.

Smith, N. A., & Trainor, L. J. (2008). Infant-Directed Speech Is Modulated by Infant Feedback. *Infancy*, *13*(4), 410–420. https://doi.org/10.1080/15250000802188719

Snow, C. E., Eeden, R., & Muysken, P. (1981). The Interactional Origins of Foreigner Talk: Municipal Employees and Foreign Workers. *International Journal of the Sociology of Language*, *1981*(28). https://doi.org/10.1515/ijsl.1981.28.81

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532. https://doi.org/10.1016/j.dr.2007.06.002

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, *128*(1), 389–400. https://doi.org/10.1121/1.3419786

Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, *44*, 1–18. https://doi.org/10.1016/j.dr.2016.12.001

Starkweather, J. A. (1967). Vocal behavior as an information channel of speaker status. *Academic Press, New York*, 253–262.

Stent, A. J., Huffman, M. K., & Brennan, S. E. (2008). Adapting speaking after evidence of misrecognition: Local and global hyperarticulation. *Speech Communication*, *50*(3), 163–178. https://doi.org/10.1016/j.specom.2007.07.005

Stewart, Bouchard-Ryan, M. A., Ellen. (1982). Attitudes Toward Younger and Older Adult Speakers: Effects of Varying Speech Rates. *Journal of Language and Social Psychology*, *1*(2), 91–109. https://doi.org/10.1177/0261927X8200100201

Suffill, E., Kutasi, T., Pickering, M., & Branigan, H. (2021). Lexical alignment is affected by addressee but not speaker nativeness. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728921000092

Tarone, E. (1980). Communication strategies, foreign talk, and repair in interlanguage. *Language Learning*, *30*(2), 417–428. https://doi.org/10.1111/j.1467-1770.1980.tb00326.x

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, *9*(2), 335–340. https://doi.org/10.3758/BF03196290

Tursunov, A., Kwon, S., & Pang, H.-S. (2019). Discriminating Emotions in the Valence Dimension from Speech Using Timbre Features. *Applied Sciences*, *9*(12), 2470. https://doi.org/10.3390/app9122470

Tweissi, A. I. (1990). Foreigner Talk in Arabic: Evidence for the universality of language simplification. *Annual Symposium on Arabic Linguistics*. Perspectives on Arabic Linguistics: Papers from the Annual Symposium on Arabic Linguistics. Volume II, Salt Lake City, Utah. https://doi.org/10.1075/cilt.72.16twe

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*(1), 2–7. https://doi.org/10.1016/j.specom.2006.10.003

Valdman, A. (1981). Sociolinguistic Aspects of Foreigner Talk. *International Journal of the Sociology of Language*, *1981*(28), 41–52. https://doi.org/10.1515/ijsl.1981.28.41

Warren-Leubecker, A., & Bohannon, J. N. (1982). The effects of expectation and feedback on speech to foreigners. *Journal of Psycholinguistic Research*, *11*(3), 207–215. https://doi.org/10.1007/BF01067564

Wooldridge, B. (2001). ‘Foreigner Talk’: An Important Element in Cross-Cultural Management Education and Training. *International Review of Administrative Sciences*, *67*(4), 621–634. https://doi.org/10.1177/0020852301674002

Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013). Vowel Hyperarticulation in Parrot-, Dog- and Infant-Directed Speech. *Anthrozoös*, *26*(3), 373–380. https://doi.org/10.2752/175303713X13697429463592

Xu Rattanasone, N., Burnham, D., & Reilly, R. G. (2013). Tone and vowel enhancement in Cantonese infant-directed speech at 3, 6, 9, and 12 months of age. *Journal of Phonetics*, *41*(5), 332–343. https://doi.org/10.1016/j.wocn.2013.06.001

Zangl, R., & Mills, D. L. (2007). Increased Brain Activity to Infant-Directed Speech in 6-and 13-Month-Old Infants. *Infancy*, *11*(1), 31–62. https://doi.org/10.1207/s15327078in1101_2

Zhang, Y. B., & Giles, H. (2017). Communication Accommodation Theory. In *The International Encyclopedia of Intercultural Communication* (pp. 1–14). American Cancer Society. https://doi.org/10.1002/9781118783665.ieicc0156

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the “what” to the “why” in second-language research. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied Sociolinguistics* (pp. 223–244). Cambridge University Press. https://doi.org/10.1017/CBO9780511663673.007

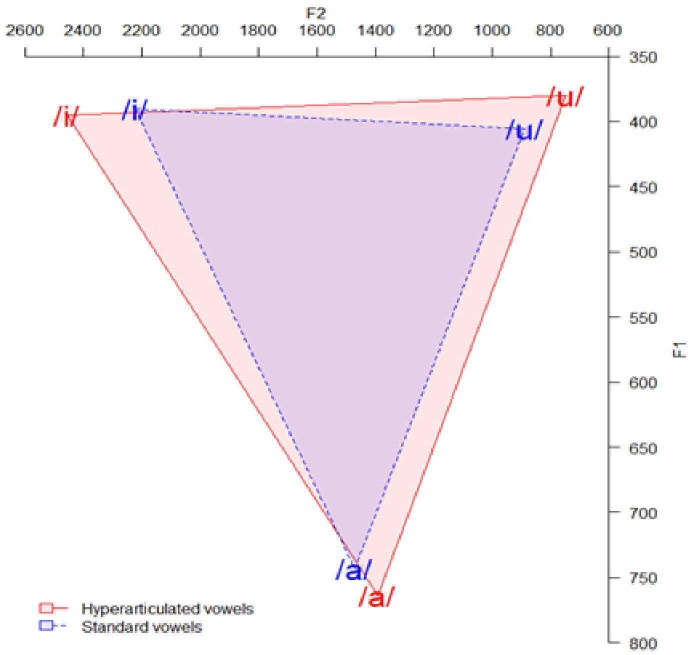

Figure 1. Example of a hyperarticulated vocalic triangle compared to the realization of standard vowels (not based on real data). The X-axis represents F2 values (Hz); the Y-axis represents F1 values (Hz).

Features of FDS	
Vowel hyperarticulation	Evidence supporting its presence
Low speech rate	
Consonant hyperarticulation	Under debate
High intensity	
Wide pitch range	
Pitch contour	

Figure 2. Summary of the features of FDS. "Evidence supporting its presence" means that there is evidence in favour of this feature. Features reporting "Under debate" mean that there is still little or mixed evidence.

Factors influencing FDS		
Listener characteristics	*Imaginary addressees*	Evidence supporting its presence
Situational and contextual factors	*Task instruction*	Evidence supporting its presence
Situational and contextual factors	*Feedback and perceived successful communication*	Evidence supporting its presence
Listener characteristics	*Language proficiency and accentedness*	Under debate
Listener characteristics	*Perception of foreignness*	Under debate
Situational and contextual factors	*Communicative goal*	Under debate
Speaker characteristics	*Previous experience and bilingualism*	Under debate
Speaker characteristics	*Emotional closeness, familiarity, and relationship*	Under debate

Figure 3. Summary of the factors influencing FDS realization. “Evidence supporting its presence” means that there is evidence in favour of this factor. Factors reporting “Under debate” mean that there is still little or mixed evidence.