

Language switching makes pronunciation less native-like

Matthew Goldrick^a, Elin Runnqvist^b, and Albert Costa^c

^aDepartment of Linguistics, Northwestern University

^bLaboratoire de Psychologie Cognitive, CNRS, Université Aix-Marseille

^cDepartamento de Tecnologia, ICREA, Universitat Pompeu Fabra

Word count: 2036

Address for correspondence:

Matthew Goldrick

Department of Linguistics

Northwestern University

2016 Sheridan Rd.

Evanston, IL 60208 USA

Phone: (847) 491-8053

Fax: (847) 491-3770

Email: matt-goldrick@northwestern.edu

Abstract

It is well known that multilinguals' non-native productions are accented. Do these deviations from monolingual productions simply reflect the mis-learning of non-native sound categories or can difficulties in processing speech sounds also contribute to accent? This is predicted by interactive theories of production, which propose that non-target representations, partially activated during lexical access, influence phonetic processing. We examine this using language switching, a task that is well known to disrupt multilingual speech production. We find that these disruptions extend to the articulation of individual speech sounds. When Spanish native speakers are required to unexpectedly switch the language of production between Spanish and English, their speech becomes more accented relative to trials where they do not switch languages (particularly for cognate targets). This suggests accents reflect not only difficulty in acquiring second language speech sounds, but also the influence of representations partially activated during on-line speech processing.

Language switching makes pronunciation less native-like

The non-native productions of multilinguals are accented; relative to monolinguals, their productions deviate towards the phonetic properties of their first language. Such deviations reflect, in part, the mis-learning of non-native sound categories (Flege, 1987). Can difficulties in speech sound processing also contribute to accent? We examine this using language switching. While proficient multilingual speakers can readily switch between languages—producing language-appropriate words (*casa* vs. *house*) and speech sounds (pronouncing the first vowel in *taxi* as /a/ vs. /æ/)—switching is challenging, resulting in disruptions to speech production (Meuter & Allport, 1999). Do these disruptions extend to phonetic processing, causing a bilingual’s native language to further contaminate non-native productions?

Such effects are predicted by interactive theories of language production. Disruptions to lexical access decrease target activation and/or increase the activation of competitors. In interactive theories, these on-line reductions in relative target activation disrupt phonetic processing on a trial-by-trial basis—a prediction confirmed by studies of monolinguals (Goldrick & Chu, in press; Kello, Plaut, & MacWhinney, 2000). Do disruptions due to language switching lead to similar effects? These theories predict that trial-to-trial variations in the degree to which native language representations are activated should influence the degree to which non-native productions are accented.

Previous work has shown the degree to which native language properties intrude into non-native productions varies across contexts. Accents increase when

speakers mix languages (e.g., reading sentences with Greek and English words) compared to single language contexts (e.g., entirely English sentences; Antoniou, Best, Tyler, & Kroos, 2011; Bullock, Toribio, González & Dalola, 2006; Flege, 1991). However, it is unclear if accents can also shift within mixed language contexts. When mixed-language sentences are read aloud, one study found accents are stronger at points where speakers plan to switch languages (Bullock et al., 2006); however, two others failed to show systematic effects (Grosjean & Miller, 1994; López, 2012).

We extend this work by examining how trial-specific processing disruptions due to unplanned switching influence accent (see also Olson, 2013, discussed below). Native Spanish speakers named pictures with a colored frame indicating Spanish or English as the response language (Meuter & Allport, 1999). If switching modulates accent, we expect participants to be more accented on switch trials (where the language of production differs from the previous trial) versus stay trials (where the language of production is the same).

We indexed accent using the contrast between voiced (/d/) and voiceless (/t/) sounds. Acoustically, these are distinguished by voice onset time (VOT; the time between the release of the consonant's constriction and the onset of periodicity signaling modal vocal fold vibration). In Spanish, vocal fold vibration typically starts before the release of voiced stops' constrictions; voiceless consonants are produced with a short positive lag between constriction release and vocal fold vibration. In contrast, in English, the voiced versus voiceless contrast is realized by a short and long positive lag respectively (Lisker & Abramson, 1964). The conflicting realizations of this contrast are reflected in native Spanish speakers' English

productions. These speakers deviate towards the phonetic properties of Spanish, producing shorter and more prevoiced VOTs than English monolinguals (Flege, 1991). We examine whether this effect is enhanced by the demands of unexpectedly switching languages.

Method

Participants

Ten native Spanish speakers from Barcelona participated. All speakers had some knowledge of Catalan (which has the same voicing contrast as Spanish; Recasens & Mira, 2013). These speakers began learning English in childhood (mean age of onset of English learning: 6.8 years; range 3-11). Proficiency scores in the three languages (reported in Table 1) were obtained through a questionnaire filled out by the participants after the experiment (see also Runnqvist & Costa, 2012; Runnqvist, Strijkers, Alario, & Costa, 2012). The scores are on a 4-point scale, where 4 = native-speaker level; 3 = advanced level; 2 = medium level; and 1 = low level of proficiency. The self-assessment index represents the average of the participants' responses in four domains (speech comprehension, speech production, reading, and writing).

Table 1. Participant age and self-rated proficiency measures (1 = low proficiency; 4 = native-speaker proficiency).

Participant	Age	Self-Rated Proficiency		
		Spanish	Catalan	English
1	19	4	4	4
2	21	4	4	3.8
3	25	4	2.6	3.2
4	30	4	4	2.67
5	20	4	4	3
6	20	4	4	3
7	25	4	3	4
8	19	4	4	4
9	22	4	4	3.2
10	19	4	4	3.8

Materials and Procedure

The 32 target items were evenly split between voiced (/d/; *door, dinero*) and voiceless (/t/; *tent, toro*) initial words. Within each language, words were evenly split between cognate and non-cognate targets and evenly distributed across four vowels (/ε/, /i/, /o/ and /α/ in Spanish; /ε/, /ɪ/, /oʊ/ and /æ/ in English; the Appendix provides a full list of stimuli). A set of 24 additional pictures served as fillers.

Pictures were presented in short sequences ranging in length unpredictably from 5 to 14 trials. Each sequence contained at least one switch between languages. For experimental trials the response language was either the same (stay) or different (switch) than the previous trial. Additionally, the target name could have the same or different initial phoneme to the previous trial. Each of these 16 targets for each language was presented twice in each of these 4 contexts (switch/stay X related/ unrelated). This yielded a total of 128 trials in each language; these 256 experimental trials were embedded within a total of 832 trials.

Before the experiment proper, participants were familiarized with the picture names. Each pictures was presented with its English and Spanish labels for 4 seconds. Participants then received both written and oral instructions and were asked to name each picture as rapidly and accurately as possible. Each trial consisted of the presentation of a blank screen (300 ms), a fixation mark (300 ms) another blank screen (300 ms) and the picture with a colored frame indicating the response language (1750 ms). Pictures disappeared from the computer screen upon detection of a response. The experiment was administered on PCs running DMDX (Forster & Forster, 2003) which also recorded the vocal responses.

Acoustic Analysis

Errors (identified by a proficient Spanish/English bilingual) and trials with recording errors (6.8%; N = 2560) were excluded. Two naïve coders performed acoustic analysis. Voice onset time (VOT) was defined as time from the burst to the onset of periodicity. Prevoicing was defined as cases where periodicity began prior to the burst. Recordings were randomly assigned to the coders. To assess reliability,

99 tokens from across several talkers were selected. VOTs on these tokens were strongly positively correlated ($r(97) = 0.87$).

Results

Linear mixed effects models including the maximal (uncorrelated) random effects structure were used to analyze the likelihood that voiced stops would be produced with a short-lag VOT and the VOT of voiceless stops¹. Significance of fixed effects was assessed using model comparison.

For voiced consonants (Figure 1A), participants successfully code-switched, producing more English-appropriate short-lag VOTs on English vs. Spanish trials (main effect of language: 95% CI [14%, 22%]; $\beta = 1.16$, s.e. = 0.25, $\chi^2(1) = 12.83$, $p < .001$). Although there was no main effect of trial type ($\chi^2(1) = 2.09$, $p < .15$), trial type and language interacted ($\beta = -0.77$, s.e. = 0.24, $\chi^2(1) = 3.75$, $p < .053$).

Participants' English productions were significantly more accented on switch versus stay trials (95% CI = [-6%, -14%]; $\chi^2(1) = 6.04$, $p < .02$), but no effect was observed for native-language Spanish productions ($\chi^2(1) < 1$).

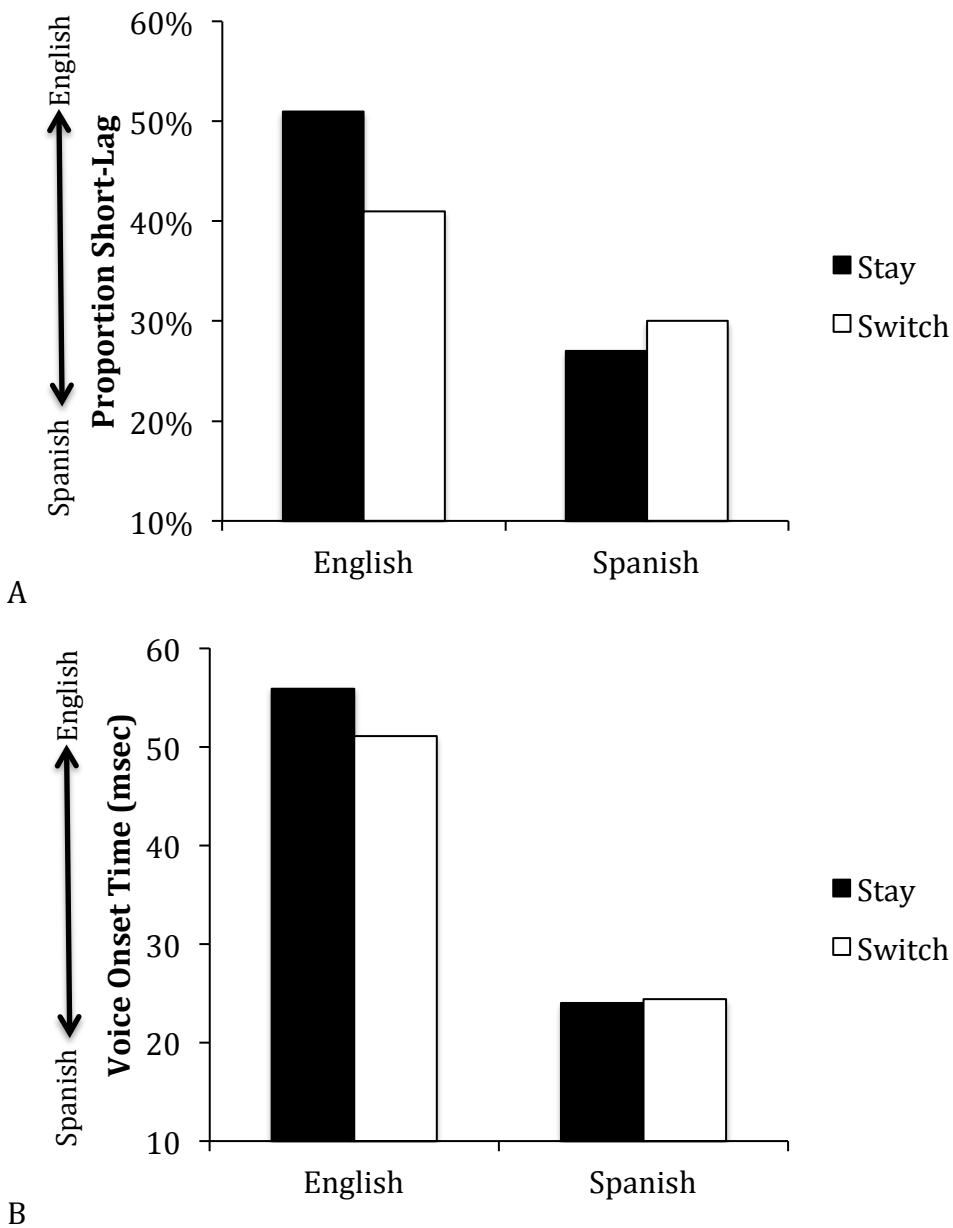
Similar results were found for voiceless consonants (Figure 1B). Participants successfully code-switched (main effect of language: 95% CI [24.3 msec, 34.2 msec]; $\beta = 27.2$, s.e. = 4.7, $\chi^2(1) = 14.66$, $p < .001$). Although there was no main effect of trial type ($\chi^2(1) = 2.33$, $p < .15$), trial type and language interacted ($\beta = -4.1$, s.e. = 1.5, $\chi^2(1) = 5.84$, $p < .02$). Participants' English productions were significantly more

¹ Similar results were observed for log-transformed positive VOTs, which reduced the non-normality of model residuals.

accented on switch versus stay trials (95% CI = [-3.7 msec, -8.6 msec]; $\chi^2(1) = 4.51$, $p < .04$) but no effect was observed in Spanish ($\chi^2(1) < 1$).

Figure 1. Phonetic realization of voiced and voiceless consonants across trial types.

A: Voiced consonants: Mean proportion of consonants realized with short-lag voice onset time (prototypical English pronunciation) versus prevoicing (prototypical Spanish pronunciation). B: Voiceless consonants: Mean voice onset time. Spanish voiceless stops are prototypically realized with relatively short voice onset times compared to English stops.

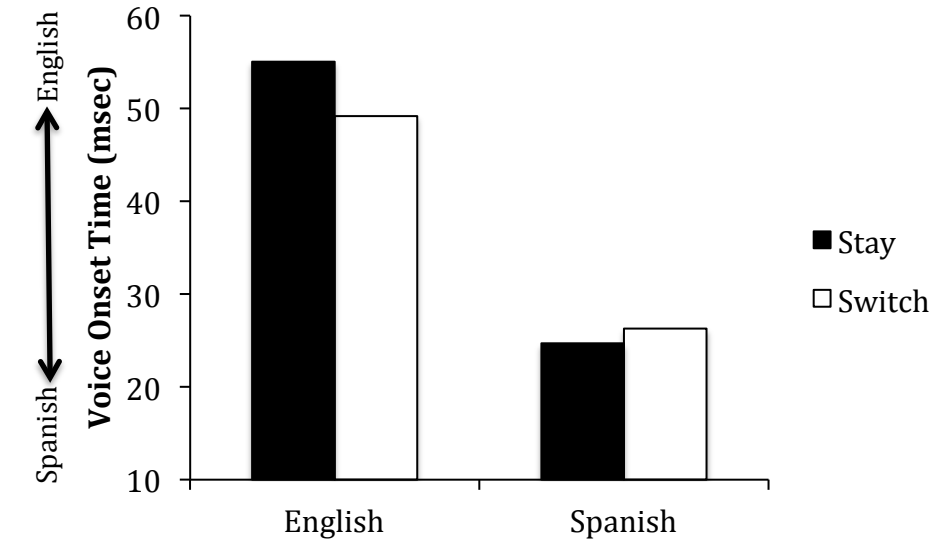


An additional set of regressions examined if the critical interaction between target type (stay/switch) and language (English/Spanish) was modulated by the phonological relatedness of the preceding trials and/or cognate status of the target (both contrast-coded).

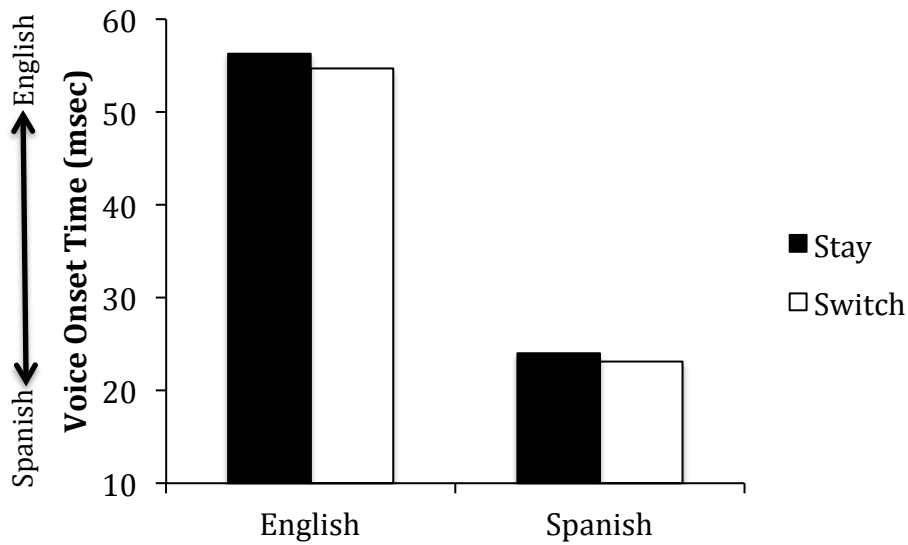
For voiced targets, the maximal, uncorrelated random effects structure that converged included only random slopes for each main effect (no interactions). Neither 3-way interaction reached significance ($\chi^2(1) < 1$).

For voiceless targets, the maximal, uncorrelated random effects structure that converged included only random slopes for each main effect, all two-way interactions, and the 3-way interaction of cognate status, trial type, and language. The 3-way interaction of trial type, language and relatedness failed to reach significance ($\chi^2(1) < 1$). In contrast, cognate status significantly modulated the interaction of trial type and language ($\beta = -7.2$, s.e. = 2.9, $\chi^2(1) = 5.69$, $p < .02$). Follow-up regressions revealed a significant interaction of trial type and language for cognate targets ($\chi^2(1) = 8.69$, $p < .005$) but not non-cognate targets ($\chi^2(1) < 1$). As shown in Figure 2, participants' English productions were more accented on switch vs. stay trials for cognates (95% CI = [-1.0 msec, -10.7 msec]) but no significant effect was observed for non-cognates (95% CI = [1.8 msec, -4.9 msec]).

Figure 2. Phonetic realization of voiceless consonants across trial types. (A) Cognate targets (*taxi*). (B) Non-cognate targets (*tenedor* ‘fork’).



A.



B.

Discussion

When required to unexpectedly switch the language of production, native Spanish speakers' English productions became more accented relative to trials

where they did not switch languages. This extends previous work suggesting that placing speakers in a difficult production context can increase accent (Gustafson, Engstler, & Goldrick, in press; Howell & Dworzynski, 2001). The current findings extend these results, showing that speakers do not simply adopt a particular speech style in a difficult production context. Within a difficult context, processing disruptions on specific trials increase accent.

These findings are consistent with a recent cued switching study (Olson, 2013) which showed Spanish-English bilinguals in the United States have increased accents on voiceless stops on stay vs. switch trials. Our results show that switching affects not only voiceless but also voiced sounds, providing clear evidence that this effect decreases the contrast between these two sound categories. Note that Olson (2013) found that this effect was found in the dominant, but not non-dominant language, and occurred only when the production context was heavily biased towards the non-dominant language (not in a balanced context). In contrast, we find, in a balanced production context, effects in the non-dominant language (which tends to be more affected by cross-language interactions across a variety of production contexts).

As noted in the introduction, studies of monolingual speakers have shown that trial-specific disruptions to lexical access can lead to disruptions to phonetic processing. To account for such effects, interactive theories have proposed that variations in the relative activation of target vs. non-target representations within lexical access can influence phonetic processing (Goldrick & Chu, in press; Kello et al., 2000). Such theories provide a ready account of the effects we observe here. On

switch trials, non-target language representations are more active than on stay trials. Interactive mechanisms—e.g., cascading activation from non-selected representations—allow these partially activated non-target representations to influence phonetic processing. This causes productions to deviate away from the target language, towards the non-target language—increasing speakers' accents.

This mechanism provides a ready account of stronger effects for voiceless stops in cognate vs. non-cognate targets. A number of studies have suggested that native language representations are activated during non-native lexical access—particularly for cognate targets (Costa, Caramazza, & Sebastián-Gallés, 2000). The greater activation of non-target language representations for cognates will cascade to phonetic processes, enhancing the degree to which phonetic properties of the non-target language intrude during production. Consistent with this account, Amengual (2012) reports that when Spanish-English bilinguals read sentences aloud, Spanish voiceless stops are more accented in cognate vs. non-cognate words.

While accents partially reflect the difficulties multilinguals have in successfully acquiring the sound system of a non-native language, this body of work shows that difficulties in language processing also play an important role. These difficulties reflect general principles of the language production system, shared by both monolingual and multilingual processing—interactive mechanisms that allow partially activated lexical and phonological representations to influence phonetic processing.

Authorship

All authors contributed to the study design. Testing and data collection were performed by E.R. under the supervision of A.C. M.G. performed the data analysis; all authors contributed to interpretation. M.G. drafted the paper. All authors approved the final version of the paper for submission.

Author Note

Supported by grants BCS0846147 from the US National Science Foundation, PSI2011-23033 and CONSOLIDER-INGENIO2010 CSD2007-00048 from the Spanish Government, and SGR 2009-1521 from the Catalan Government.

References

- Amengual, M. (2012). Influence bilingual in interlingual speech: cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, 15, 517-530.
- Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, 39, 558-570.
- Bullock, B. E., Toribio, A.J., González, V., & Dalola, A (2006). Language dominance and performance outcomes in bilingual pronunciation. In M. G. O'Brien, C. Shea, & J. Archibald J (Eds.) *Proceedings of the 8th Generative Approaches to Second Language Acquisition (GASLA): The Banff conference* (pp. 9-16). Somerville, MD: Cascadilla Press.
- Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1283-1296.
- Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.
- Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89, 395-411.
- Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15, 67-83.

- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers*, *35*, 116–124.
- Goldrick, M., & Chu, K. (in press). Gradient co-activation and speech error articulation: Comment on Pouplier and Goldstein (2010). *Language, Cognition and Neuroscience*.
- Grosjean, F., & Miller, J. L. (1994). Going in and out of languages: An example of bilingual flexibility. *Psychological Science*, *5*, 201-209.
- Gustafson, E., Engstler, C., & Goldrick, M. (in press). Phonetic processing of non-native speech in semantic vs. non-semantic tasks. *Journal of the Acoustical Society of America*.
- Howell, P., & Dworzynski, K. (2001). Strength of German accent under altered auditory feedback. *Perception & Psychophysics*, *63*, 501-513.
- Kello, C. T., Plaut, D. C., & MacWhinney, B. (2000). The task dependence of staged versus cascaded processing: An empirical and computational study of Stroop interference in speech production. *Journal of Experimental Psychology: General*, *129*, 340–360.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*, 384-422.
- López, V. G. (2012). Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Language Research*, *28*, 243-263.

- Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25-40.
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, *41*, 407-420.
- Recasens, D., & Mira, M. (2013). Voicing assimilation in Catalan three-consonant clusters. *Journal of Phonetics*, *41*, 264-280.
- Runnqvist, E. & Costa, A. (2012). Is retrieval-induced forgetting behind the bilingual disadvantage in speech production? *Bilingualism: Language and Cognition*, *15*, 365-377.
- Runnqvist, E., Strijkers, K., Alario, F-X., & Costa, A. (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language*, *66*, 850-869.

Appendix

Spanish stimulus	English translation	English stimulus
tenedor	<i>fork</i>	tent
delantal	<i>apron</i>	desk
tijeras	<i>scissors</i>	tin
dinero	<i>money</i>	dish
toro	<i>bull</i>	tor
doce	<i>twelve</i>	door
taza	<i>cup</i>	tap
dado	<i>dice</i>	dancer
teléfono	<i>telephone</i>	telephone
dentista	<i>dentist</i>	dentist
tinte	<i>tint</i>	tint
diploma	<i>diploma</i>	diploma
tornado	<i>tornado</i>	tornado
dormitorio	<i>dormitory</i>	dormitory
taxi	<i>taxi</i>	taxi
dálmata	<i>Dalmatian</i>	Dalmatian